



Τεχνολογικό Εκπαιδευτικό Ίδρυμα Κρήτης

**Σχολή Τεχνολογικών Εφαρμογών
Τμήμα Εφαρμοσμένης Πληροφορικής και Πολυμέσων**



Πτυχιακή εργασία

**Ανίχνευση κίνησης από κάμερες για
λόγους ασφαλείας**

Σταματάκης Ξενοφών (Α.Μ. 1878)

Επιβλέπων Καθηγητής: Τριανταφυλλίδης Γεώργιος

Abstract

Motion detection is a significant field of computer vision both as an object of research as well as an application. It is realized through the automated recording and analysis of data captured by camera. In our days, there is a significant number of applications using this kind of software, both in the field of security services, for example surveillance cameras, as well as in cases where human-computer interaction is necessary. In addition, it is used in patient monitoring systems and for monitoring people with disabilities (fall detection system).

In the present dissertation, to begin with, we define the concept of camera and examine the types of security cameras and their uses. We evaluate the most widespread types of vision in cameras used for these purposes and introduce their advantages and disadvantages. Furthermore, we make an analysis of the concept of human detection by means of security cameras, and demonstrate the various methods of implementation: face detection, movement detection, skin detection and point feature tracking. Finally, we develop an application which aims to analyze movement and detect potential human presence. We introduce the program by demonstrating screenshots from various occasions. The application is tested in indoor and outdoor spaces, with natural or artificial light, in the insertion of a new object in the shot, in dark spaces and with light alternations.

Σύνοψη

Η ανίχνευση κίνησης είναι ένας σημαντικός τομέας της τεχνητής όρασης τόσο σαν αντικείμενο έρευνας όσο και σε επίπεδο εφαρμογών. Πραγματοποιείται μέσα από την αυτοματοποιημένη καταχώρηση και ανάλυση δεδομένων που έχουν καταγράψει από κάμερα. Στις μέρες μας, υπάρχει σημαντικός αριθμός εφαρμογών που χρησιμοποιούν τέτοιου είδους λογισμικό, τόσο στον τομέα της ασφάλειας, π.χ. συστήματα παρακολούθησης, όσο και σε τομείς που απαιτείται αλληλεπίδραση μεταξύ ανθρώπου και υπολογιστή. Ακόμη χρησιμοποιείται και σε συστήματα παρακολούθησης ασθενών, καθώς και για την παρακολούθηση ατόμων με αναπηρία (σύστημα ανίχνευσης πτώσης).

Στην παρούσα πτυχιακή ξεκινώντας, ορίζουμε την έννοια της κάμερας και εξετάζουμε τα ειδή κάμερας ασφάλειας και τις χρήσεις τους. Αξιολογούμε τα πιο διαδεδομένα ειδή λήψης στις κάμερες που χρησιμοποιούνται για αυτό το σκοπό και παρουσιάζουμε τα πλεονεκτήματα και τα μειονεκτήματά τους. Στη συνέχεια, αναλύουμε την ανίχνευση ανθρώπου, με τη βοήθεια κάμερας ασφάλειας, και παρουσιάζουμε τις διαφορές μεθόδους υλοποίησης της: ανίχνευση προσώπου, ανίχνευση κίνησης, ανίχνευση δέρματος και ο εντοπισμός χαρακτηριστικών σημείων. Τέλος, αναπτύσσουμε μια εφαρμογή της οποίας στόχος είναι, με καταγεγραμμένο υλικό από κάμερα, να αναλύσει την εκάστου κίνηση και να ανιχνεύσει πιθανή ανθρώπινη παρουσία. Παρουσιάσουμε το πρόγραμμα επιδεικνύοντας στιγμιότυπα από διαφορετικές περιπτώσεις. Το πρόγραμμα δοκιμάζεται σε εσωτερικό και εξωτερικό χώρο, με φυσικό ή τεχνητό φως, στην εισαγωγή αντικείμενου στο χώρο, σε σκοτεινό περιβάλλον και στην εναλλαγή φωτισμού.

Περιεχόμενα

Abstract.....	3
Σύνοψη	4
1 Εισαγωγή	8
1.1 Περίληψη	8
1.2 Κίνητρο.....	8
1.3 Σκοπός και Στόχος Εργασίας.....	9
1.4 Δομή εργασίας.....	9
2 Μέθοδος Ανάλυσης και Ανάπτυξης της Πτυχιακής.....	10
2.1 Γενικά	10
2.2 Θεωρίες.....	10
2.3 Τεχνολογίες.....	10
3 Εισαγωγή στις Κάμερες.....	12
3.1 Είδη κάμερας	13
3.2 Κάμερες και νυχτερινή λήψη.....	14
3.3 Πλεονεκτήματα-Μειονεκτήματα	14
4 Τεχνικές Ανίχνευσης Ανθρώπου	17
4.1 Ανίχνευση Προσώπου.....	17
4.1.1 Information based (Top-Down).....	18
4.1.2 Feature Invariant (Bottom-up).....	18
4.1.3 Template matching methods.....	19
4.1.4 Appearance based.....	19
4.2 Ανίχνευση Κίνησης.....	21
4.3 Ανίχνευση Δέρματος.....	23
4.3.1 Ανίχνευση δέρματος και χώροι χρωμάτων.....	25
4.3.2 Ταξινομητές δέρματος.....	31
4.4 Εντοπισμός χαρακτηριστικών σημείων	32
4.4.1 Κατανόηση του προβλήματος	33
4.4.2 Κατάτμηση εικόνας.....	34
4.4.3 Εντοπισμός χαρακτηριστικών σημείων.....	35
4.4.4 Ανίχνευση Μοντέλου Κίνησης.....	36
5 Υλοποίηση	38
5.1 Εισαγωγή στο Matlab	38
5.2 Εισαγωγή	45
5.3 Εφαρμογή.....	46
6 Αποτελέσματα.....	71
6.1 Ανασκόπηση.....	71
6.2 Συμπεράσματα.....	71
Βιβλιογραφία.....	72

Πίνακας Εικόνων

Εικόνα 1: Ο John Logie Baird, με την πρώτη βιντεοκάμερα	12
Εικόνα 2: Κάμερες Ασφαλείας.....	13
Εικόνα 3: Παράδειγμα low light απεικόνισης.....	14
Εικόνα 4: Παράδειγμα θερμικής κάμερας.....	15
Εικόνα 5: Παράδειγμα κάμερας «Near Infrared Illumination»	15
Εικόνα 6: Εικόνα που παρουσιάζει την ανίχνευση προσώπου σε εξωτερικό χώρο	17
Εικόνα 7: Παράδειγμα 1 από Eigenfaces	20
Εικόνα 8: Παράδειγμα 2 από Eigenfaces	20
Εικόνα 9: Παράδειγμα αλγορίθμου ανίχνευσης κίνησης	21
Εικόνα 10: Ανίχνευση Δέρματος.....	23
Εικόνα 11: Διάγραμμα πυκνότητας δέρματος Ασιατών σε διαφορετικούς χρωματικούς χώρους	25
Εικόνα 12: Διάγραμμα πυκνότητας δέρματος Ασιατών, Αφρικανών και Καυκάσιων σε διαφορετικούς χρωματικούς χώρους.	26
Εικόνα 13: RGB χρωματικός χώρος.....	27
Εικόνα 14: Αναπαράσταση του $YCbCr$ χρωματικού χώρου.....	28
Εικόνα 15: Αναπαράσταση του YIQ χρωματικού χώρου	28
Εικόνα 16: HSL και HSV.....	29
Εικόνα 17: Αναπαράσταση του CIE-lab.	30
Εικόνα 18: Ακολουθία από σημεία του Johansson που δείχνει την πλαϊνή λήψη ενός ανθρώπου που περπατάει.....	32
Εικόνα 19: Κατάτμηση Εικόνας.....	33
Εικόνα 20: Matlab Desktop (Επιφάνεια Εργασίας Matlab)	39
Εικόνα 21: Παράθυρο Εντολών (Command Window).....	39
Εικόνα 22: Ιστορικό Εντολών (Command History)	40
Εικόνα 23: Χώρος εργασίας(Workspace).....	40
Εικόνα 24: Τρέχων Κατάλογος (Current Directory)	41
Εικόνα 25: Παράθυρο Εισαγωγής Κειμένου (Editor window).....	42
Εικόνα 26: Παράθυρο Σχημάτων (Figure window)	43
Εικόνα 27: Παράθυρο Βοήθειας (Help Window).....	43
Εικόνα 28: Παράδειγμα χρήσης του Matlab στην επεξεργασία εικόνας.....	44
Εικόνα 29: Γραφική Αναπαράσταση Αλγορίθμου	45
Εικόνα 30: Παράθυρο επιλογής βίντεο	47
Εικόνα 31: Επιλογή αρχείου βίντεο προς επεξεργασία	48
Εικόνα 32: Στιγμιότυπο 1 από την αναπαραγωγή βίντεο1	48
Εικόνα 33: Στιγμιότυπο 2 από την αναπαραγωγή βίντεο 1	49
Εικόνα 34: Στιγμιότυπο 3 από την αναπαραγωγή βίντεο 1	49
Εικόνα 35: Στιγμιότυπο από το περιβάλλον εργασίας του Matlab.....	50
Εικόνα 36: Μήνυμα ανίχνευσης ανθρώπου στο βίντεο 1.....	50
Εικόνα 37: Η εικόνα που εντοπίστηκε ανθρώπινη παρουσία με και χωρίς αφαίρεση background στο βίντεο 1	51
Εικόνα 38: Στιγμιότυπο από το περιβάλλον εργασίας του Matlab μετά την εκτέλεση της εφαρμογής για το βίντεο 1	51
Εικόνα 39: Στιγμιότυπο 1 από την αναπαραγωγή βίντεο 2	52
Εικόνα 40: Στιγμιότυπο 2 από την αναπαραγωγή βίντεο 2	52
Εικόνα 41: Στιγμιότυπο 3 από την αναπαραγωγή βίντεο 2	53
Εικόνα 42: Μήνυμα ανίχνευσης ανθρώπου στο βίντεο 2.....	53
Εικόνα 43: Η εικόνα που εντοπίστηκε ανθρώπινη παρουσία στο βίντεο 2 με και χωρίς αφαίρεση background.....	54
Εικόνα 44: Στιγμιότυπο 1 από την αναπαραγωγή βίντεο 3	54
Εικόνα 45: Στιγμιότυπο 2 από την αναπαραγωγή βίντεο 3	55
Εικόνα 46: Στιγμιότυπο 3 από την αναπαραγωγή βίντεο 3	55

Εικόνα 47: Μήνυμα ανίχνευσης ανθρώπου στο βίντεο 3.....	55
Εικόνα 48: Η εικόνα που εντοπίστηκε ανθρώπινη παρουσία στο βίντεο 3 με και χωρίς αφαίρεση background.....	56
Εικόνα 49: Στιγμιότυπο από το περιβάλλον εργασίας του Matlab μετά την εκτέλεση της εφαρμογής για το βίντεο 3.....	57
Εικόνα 50: Στιγμιότυπο 1 από την αναπαραγωγή βίντεο 4.....	57
Εικόνα 51: Στιγμιότυπο 2 από την αναπαραγωγή βίντεο 4.....	58
Εικόνα 52: Στιγμιότυπο 3 από την αναπαραγωγή βίντεο 4.....	58
Εικόνα 53: Στιγμιότυπο 4 από την αναπαραγωγή βίντεο 4.....	59
Εικόνα 54: Στιγμιότυπο 5 από την αναπαραγωγή βίντεο 4.....	59
Εικόνα 55: Μήνυμα ανίχνευσης ανθρώπου στο βίντεο 4.....	59
Εικόνα 56: Η εικόνα που εντοπίστηκε ανθρώπινη παρουσία στο βίντεο 4 με και χωρίς αφαίρεση background.....	60
Εικόνα 57: Εναλλαγή αρχικής εικόνας όταν αλλάζει ο φωτισμός.....	60
Εικόνα 58: Στιγμιότυπο 1 από την αναπαραγωγή βίντεο 5.....	61
Εικόνα 59: Στιγμιότυπο 2 από την αναπαραγωγή βίντεο 5.....	61
Εικόνα 60: Στιγμιότυπο 3 από την αναπαραγωγή βίντεο 5.....	62
Εικόνα 61: Στιγμιότυπο 4 από την αναπαραγωγή βίντεο 5.....	62
Εικόνα 62: Στιγμιότυπο 4 από την αναπαραγωγή βίντεο 5.....	63
Εικόνα 63: Μήνυμα ανίχνευσης ανθρώπου στο βίντεο 5.....	63
Εικόνα 64: Η εικόνα που εντοπίστηκε ανθρώπινη παρουσία στο βίντεο 5 με και χωρίς αφαίρεση background.....	63
Εικόνα 65: Εντοπισμός pixel στο χρώμα του δέρματος.....	64
Εικόνα 66: Στιγμιότυπο 1 από την αναπαραγωγή βίντεο 6.....	65
Εικόνα 67: Στιγμιότυπο 2 από την αναπαραγωγή βίντεο 6.....	65
Εικόνα 68: Στιγμιότυπο 3 από την αναπαραγωγή βίντεο 6.....	66
Εικόνα 69: Στιγμιότυπο 4 από την αναπαραγωγή βίντεο 6.....	66
Εικόνα 70: Μήνυμα ανίχνευσης ανθρώπου στο βίντεο 6.....	67
Εικόνα 71: Η εικόνα που εντοπίστηκε ανθρώπινη παρουσία στο βίντεο 6 με και χωρίς αφαίρεση background.....	67
Εικόνα 72: Εντοπισμός pixel στο χρώμα του δέρματος για το βίντεο 6.....	68
Εικόνα 73: Στιγμιότυπο 1 από την αναπαραγωγή βίντεο 7.....	68
Εικόνα 74: Στιγμιότυπο 2 από την αναπαραγωγή βίντεο 7.....	69
Εικόνα 75: Στιγμιότυπο 3 από την αναπαραγωγή βίντεο 7.....	69
Εικόνα 76: Μήνυμα ανίχνευσης ανθρώπου στο βίντεο 7.....	70
Εικόνα 77: Εικόνα επιστροφής εφαρμογής σε περίπτωση που δεν εντοπιστεί άνθρωπος.....	70

1 Εισαγωγή

1.1 Περίληψη

Η ανίχνευση κίνησης είναι ένας σημαντικός τομέας της τεχνητής όρασης τόσο σαν αντικείμενο έρευνας όσο και σε επίπεδο εφαρμογών. Η υλοποίηση της γίνεται μέσα από την αυτοματοποιημένη καταχώρηση και ανάλυση δεδομένων που έχουν καταγράψει από κάμερα. Στις μέρες μας, υπάρχει σημαντικός αριθμός εφαρμογών που χρησιμοποιούν τέτοιου είδους λογισμικό, τόσο στον τομέα της ασφάλειας, π.χ. συστήματα παρακολούθησης, όσο και σε τομείς που απαιτείται αλληλεπίδραση μεταξύ ανθρώπου και υπολογιστή. Ακόμη χρησιμοποιείται και σε συστήματα παρακολούθησης ασθενών, καθώς και για την παρακολούθηση ατόμων με αναπηρία (σύστημα ανίχνευσης πτώσης). Ζωτική σημασία στην διαδικασία της ανίχνευσης ανθρώπινης κίνησης παίζει γενικότερα η κάμερα και πιο συγκεκριμένα το εκάστοτε είδος κάμερας ασφάλειας.

1.2 Κίνητρο

Η ικανότητα να ανιχνεύεις με αξιοπιστία ανθρώπινη κίνηση είναι ένα χρήσιμο εργαλείο για υψηλού επιπέδου εφαρμογές οι οποίες βασίζονται σε οπτική είσοδο δεδομένων. Το να αλληλεπιδράς με ανθρώπους και να αντιλαμβάνεσαι τις κινήσεις τους είναι ο πυρήνας αρκετών προβλημάτων σε ευφυή συστήματα, όπως αλληλεπίδραση άνθρωπου-υπολογιστή και ρομποτική. Ένας αλγόριθμος για την ανίχνευση ανθρώπινης κίνησης συνοψίζει ένα υψηλού εύρους ζώνης βίντεο σε μια συμπαγή περιγραφή ανθρώπινης κίνησης στο χώρο. Μερικά παραδείγματα εφαρμογών τα οποία μπορούν να πραγματοποιηθούν με μια αξιόπιστη ανίχνευση ανθρώπου είναι:

1. Αυτοματοποιημένη παρακολούθηση για λόγους ασφαλείας σε χώρους όπως αεροδρόμια, καζίνο, μουσεία, και κυβερνητικές εγκαταστάσεις. Το λογισμικό αυτό θα μπορούσε να παρακολουθεί κάμερες ασφαλείας και να ανιχνεύει ύποπτη δραστηριότητα.
2. Ανθρώπινη αλληλεπίδραση με ρομπότ. Αυτόνομα ρομπότ στο χώρο εργασίας ή στο σπίτι θα μπορούσαν να αλληλεπιδρούν πιο ομαλά με τους ανθρώπους του περιβάλλοντος αν ήταν σε θέση να ανιχνεύουν την κίνηση τους με μεγαλύτερη αξιοπιστία. Για παράδειγμα, τα ρομπότ που βοηθούν τους ηλικιωμένους θα ήξεραν τότε χρειάζεται η βοήθεια τους βασισμένα στην κίνηση του ανθρώπου.
3. Οι συσκευές ανίχνευσης πεζών σε μηχανοκίνητα οχήματα. Το λογισμικό θα μπορούσε να λειτουργεί από μία κάμερα η οποία θα βρίσκεται στο όχημα και θα ειδοποιούσε τον οδηγό αν ανίχνευε κάποιον πεζό.
4. Αυτόματη καταγραφή κίνησης για ταινίες και τηλεόραση, η οποία θα παρήγαγε μια πιο αληθοφανή κίνηση. Κάτι τέτοιο σήμερα γίνεται με διάφορους αισθητήρες κίνησης οι οποίοι τοποθετούνται στα άκρα του ηθοποιού. Με ακριβείς αλγορίθμους για ανίχνευση ανθρώπινης κίνησης, τα ίδια δεδομένα θα μπορούσαν να εξαχθούν από οποιοδήποτε βίντεο χωρίς τη χρήση εξοπλισμού.

1.3 Σκοπός και Στόχος Εργασίας

Στην παρούσα πτυχιακή εργασία θα γίνει λόγος για τις τεχνικές ανίχνευσης ανθρώπινης κίνησης χρησιμοποιώντας απλές κάμερες με τη βοήθεια του Matlab.

Αρχικά θα γίνει μια αναφορά στις εφαρμογές αναγνώρισης ανθρώπου, οι οποίες θα υλοποιηθούν μέσω τεχνικών όπως ανίχνευσης δέρματος (skin detection). Θα συζητηθούν τα πλεονεκτήματα και τα μειονεκτήματα της κάθε τεχνικής και θα κάνουμε μια σύγκριση με σκοπό να βρούμε τον αποδοτικότερο αλγόριθμο για την κάθε περίπτωση

Στη συνέχεια θα κάνουμε μια έρευνα για τις όλες τις τεχνολογίες κάμερας που μπορούν να χρησιμοποιηθούν για τη συγκεκριμένη εφαρμογή όπως είναι οι κάμερες night vision ή οι thermal cameras και θα αποφασίσουμε ποια είναι η αποδοτικότερη.

Τέλος θα υλοποιήσουμε μια εφαρμογή σε Matlab, η οποία θα λειτουργεί για ένα συγκεκριμένο είδος κάμερας και θα ανιχνεύει ανθρώπινη κίνηση σε ένα χώρο. Το συγκεκριμένο πρόγραμμα στοχεύει στην παρακολούθηση ενός χώρου για λόγους ασφαλείας

1.4 Δομή εργασίας

Κεφάλαιο 2 - Μέθοδος ανάλυσης και ανάπτυξης της πτυχιακής: Στο κεφάλαιο αυτό αναφέρονται συνοπτικά το θέμα της πτυχιακής και οι τεχνολογίες που χρησιμοποιήθηκαν.

Κεφάλαιο 3 – Εισαγωγή στις κάμερες: Στο κεφάλαιο αυτό γίνεται μια αναδρομή στα είδη κάμερας που μπορούμε να χρησιμοποιήσουμε. Αναφέραμε τις τεχνολογίες που επιτρέπουν την λήψη σε περιβάλλοντα χωρίς φωτισμό και τέλος αναλύουμε τα πλεονεκτήματα και τα μειονεκτήματα της κάθε τεχνολογίας.

Κεφάλαιο 4 –Ανίχνευση ανθρώπου: Στο κεφάλαιο αυτό γίνεται μια ανασκόπηση των δημοφιλέστερων τεχνικών βάσει των οποίων μπορεί να επιτευχθεί ανίχνευση ανθρώπου, περιγράφουμε πως λειτουργεί η κάθε τεχνική και αξιολογούμε τις μεθόδους αυτές.

Κεφάλαιο 5 - Υλοποίηση: Στο συγκεκριμένο κεφάλαιο γίνεται ανάλυση της εφαρμογής που αναπτύξαμε. Αρχικά παρουσιάζεται το εργαλείο το οποίο χρησιμοποιήθηκε για τη δημιουργία της εφαρμογής. Στη συνέχεια γίνεται μια αναφορά στα προβλήματα που συναντήσαμε και προτείνουμε κάποιες λύσεις για την αντιμετώπιση τους. Τέλος, παρουσιάζουμε την εφαρμογή μας στα πιο χαρακτηριστικά της στιγμιότυπα.

Κεφάλαιο 7 - Συμπεράσματα: Στο τελευταίο αυτό κεφάλαιο μιλάμε για τα συμπεράσματα που βγάλαμε από τη διαδικασία υλοποίησης της πτυχιακής εργασίας.

2 Μέθοδος Ανάλυσης και Ανάπτυξης της Πτυχιακής

2.1 Γενικά

Το θέμα με το οποίο ασχολούμαστε στη συγκεκριμένη πτυχιακή εργασία είναι η κατασκευή με τη βοήθεια του Matlab, μιας εφαρμογής, για συστήματα ασφαλείας που αυτοματοποιημένα θα ανιχνεύει την κίνηση που προέρχεται από ανθρώπινη παρουσία και θα επιστρέφει ένα μήνυμα στο χρήστη προκειμένου να του γνωστοποιεί αν εντόπισε άνθρωπο σε ένα βίντεο.

Ο χρήστης θα εισάγει ένα βίντεο και αυτό στη συνέχεια θα αναλύεται προκειμένου να εξακριβωθεί αν υπάρχει ανθρώπινη παρουσία στη σκηνή. Οι τεχνικές που θα χρησιμοποιήσουμε για τη δημιουργία της εφαρμογής είναι η αφαίρεση φόντου και η ανίχνευση δέρματος.

Με τη βοήθεια του Matlab καταφέραμε να επιλύσουμε κάποια πολύ βασικά προβλήματα του συστήματος ανίχνευσης ανθρώπινης παρουσίας, όπως αυτό της εισαγωγής νέων αντικειμένων στη σκηνή και τις απότομες αλλαγές του φωτισμού. Επιτύχαμε τον στόχο μας δημιουργώντας μια απλή εφαρμογή ανίχνευσης ανθρώπου που λειτουργεί αρκετά αποδοτικά κάτω από ορισμένες προϋποθέσεις.

2.2 Θεωρίες

Θα αναφέρουμε διάφορες βασικές μεθόδους που συναντήσαμε κατά τη διάρκεια της έρευνας. Τις αναλύουμε, τις αξιολογούμε και στη συνέχεια επιλέγουμε κάποιες από αυτές για την υλοποίηση της εργασίας μας.

Αρχικά παρουσιάζεται η ανίχνευση προσώπου, καθώς είναι ένας πολύ αποτελεσματικός τρόπος για την ανίχνευση ανθρώπου. Το ανθρώπινο πρόσωπο είναι μοναδικό, και γι αυτό το λόγο αν καταφέρει να ανιχνευθεί με επιτυχία οδηγεί σε ακριβή ανίχνευση ανθρώπου.

Στη συνέχεια, αναφέρουμε την ανίχνευση κίνησης ο οποίος είναι ένας πολύ αποτελεσματικός τρόπος για να ανιχνεύουμε αλλαγές στην εικόνα, είτε αυτές προέρχονται από ενδιαφέροντα κίνηση όπως είναι αυτή του ανθρώπου είτε από παραπλανητική όπως η αλλαγή στοιχείων του φόντου.

Επειτα θα μιλήσουμε για την ανίχνευση δέρματος και τους χρωματικούς χώρους με βάση τους οποίους μπορούμε να ομαδοποιήσουμε το ανθρώπινο δέρμα, ανεξαρτήτως του φωτισμού της λήψης. Το ανθρώπινο δέρμα έχει μοναδικά χαρακτηριστικά τόσο από άποψη υφής όσο και από άποψη χρώματος και γι αυτό το λόγο ένα σύστημα βασισμένο στην ανίχνευση δέρματος έχει τη δυνατότητα να προσφέρει αξιόλογα αποτελέσματα.

Τέλος, θα παρουσιάσουμε μια μέθοδο που βασίζεται στον εντοπισμό χαρακτηριστικών σημείων. Θα ερευνήσουμε κατά πόσο η κίνηση των άκρων προσφέρει αρκετές πληροφορίες προκειμένου να εντοπιστεί ανθρώπινη παρουσία και θα μελετήσουμε την αποτελεσματικότητα ενός τέτοιου συστήματος.

2.3 Τεχνολογίες

Στο σημείο αυτό θα μιλήσουμε για τη γλώσσα προγραμματισμού που χρησιμοποιήθηκε για τη δημιουργία της εφαρμογής μας.

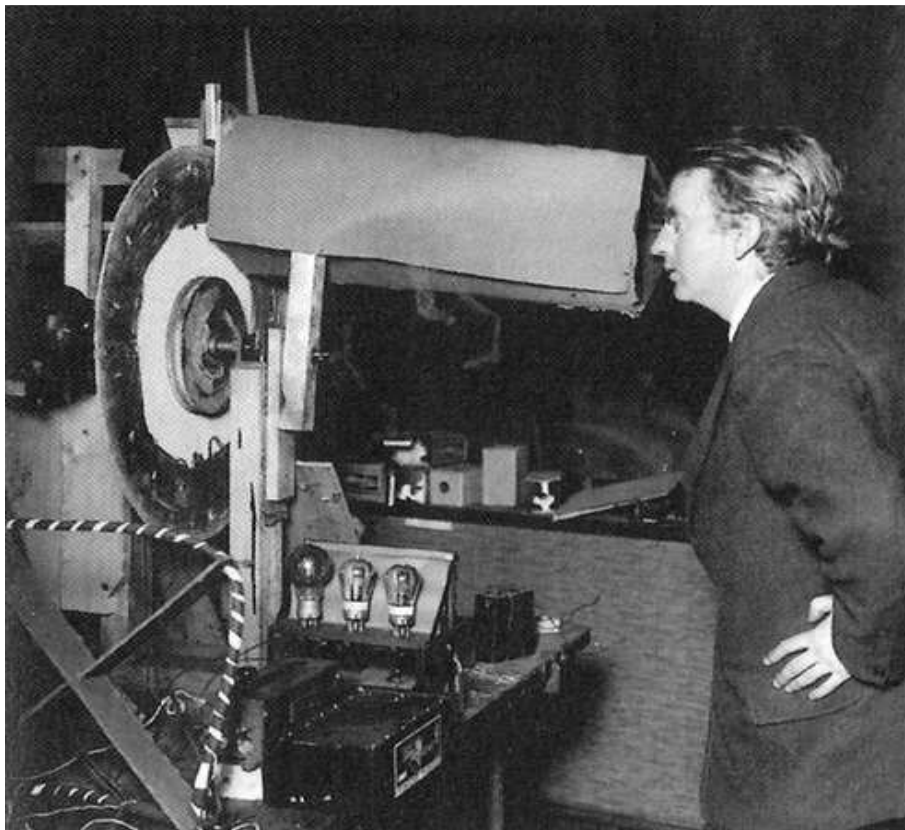
Το Matlab είναι μια υψηλού επιπέδου γλώσσα προγραμματισμού σε ένα διαδραστικό περιβάλλον, που χρησιμοποιείτε για αριθμητικούς υπολογισμούς, οπτικοποίηση δεδομένων

και προγραμματισμό. Με τη βοήθεια του Matlab μπορούμε να αναλύσουμε δεδομένα, να αναπτύξουμε αλγόριθμους, και να δημιουργήσουμε μοντέλα και εφαρμογές. Τα εργαλεία της γλώσσας αυτής και οι ενσωματωμένες μαθηματικές συναρτήσεις μας επιτρέπουν να εξερευνήσουμε πολλές προσεγγίσεις και να φτάσουμε στη λύση ενός προβλήματος πολύ πιο γρήγορα από άλλες παραδοσιακές γλώσσες προγραμματισμού όπως η C/C++ ή η Java.

Μπορούμε να χρησιμοποιήσουμε το Matlab για ένα ευρύ φάσμα εφαρμογών, συμπεριλαμβανομένων των: επεξεργασία σήματος και επικοινωνιών, συστήματα έλεγχου, επεξεργασία εικόνας και βίντεο. Μια πιο αναλυτική περιγραφή του Matlab θα ακολουθήσει σε επόμενο κεφάλαιο.

3 Εισαγωγή στις Κάμερες

Η βιντεοκάμερα είναι μια φωτογραφική μηχανή που χρησιμοποιείται για την απόκτηση εικόνας κίνησης. Αρχικά, αναπτύχθηκε από την τηλεοπτική βιομηχανία αλλά τώρα είναι πολύ διαδεδομένη και σε άλλες εφαρμογές. Οι πρώτες βιντεοκάμερες είναι αυτές του John Logie Baird, βασισμένες στον ηλεκτρομηχανικό δίσκο Nipkow και χρησιμοποιήθηκαν από το BBC, σε πειραματικές εκπομπές από το 1930. Όλα τα ηλεκτρονικά σχέδια που είχαν ως βάση τον καθοδικό σωλήνα, όπως το εικονοσκόπιο του Vladimir Zworykin και τον image dissector του Philo T. Farnsworth, αντικατέστησαν την τεχνολογία του Baird το 1940 και χρησιμοποιήθηκαν ευρέως μέχρι τα τέλη του 1980, όταν οι κάμερες με αισθητήρες εικόνας στερεάς κατάστασης όπως CCD και αργότερα οι CMOS με ενεργούς αισθητήρες εικόνας εξάλειψαν τα κοινά προβλήματα της τεχνολογίας σωλήνα, όπως το "κάψιμο" της εικόνας και έκανε την ψηφιακή ροή εργασίας πιο πρακτική. Οι βιντεοκάμερες χρησιμοποιούνται κυρίως με δυο τρόπους. Ο πρώτος, είναι στη ζωντανή τηλεόραση, όπου η κάμερα τροφοδοτεί σε πραγματικό χρόνο εικόνες απευθείας σε μια οθόνη για άμεση παρατήρηση. Λίγες κάμερες συνεχίζουν να εξυπηρετούν προϊόντα ζωντανής τηλεόρασης αφού οι περισσότερες ζωντανές συνδέσεις είναι για λόγους ασφαλείας, για στρατιωτικές και εμπορικές επιχειρήσεις στις οποίες απαιτείται ζωντανή προβολή. Στο δεύτερο τρόπο, οι εικόνες αποθηκεύονται σε μια αποθηκευτική μονάδα για περαιτέρω επεξεργασία, για πολλά χρόνια η βιντεοκασέτα ήταν η κυρίαρχη μορφή για αυτό το σκοπό, αλλά τα οπτικά μέσα αποθήκευσης την αντικατέστησαν. Καταγεγραμμένα βίντεο χρησιμοποιούνται στην τηλεόραση και στην παρακολούθηση και επιτήρηση εργασιών όπου η χωρίς παρακολούθηση καταγραφή μιας κατάστασης απαιτείται για μετέπειτα ανάλυση.



Εικόνα 1: Ο John Logie Baird, με την πρώτη βιντεοκάμερα

3.1 Είδη κάμερας

Οι κάμερες ασφαλείας μπορούν να διαχωριστούν με βάση τον τύπο τους σε τέσσερις κατηγορίες. Αρχικά έχουμε τις κάμερες dome(θόλου) οι οποίες πήραν το όνομα τους διότι βρίσκονται εντός προστατευτικού θόλου με αποτέλεσμα να είναι ανθεκτικότερες σε προσπάθεια βανδαλισμού τους. Επιπλέον ο θόλος βοηθά έτσι ώστε να μην είναι εμφανής η κατεύθυνση στην οποία είναι στραμμένη η κάμερα. Στη δεύτερη κατηγορία έχουμε τις κάμερες box οι οποίες έχουν σχήμα κουτιού και είναι οι πλέον συνήθεις σε μορφή, κάμερες. Έχουν τη δυνατότητα να τοποθετηθούν σε οποιαδήποτε κατακόρυφη επιφάνεια με τη βοήθεια κατάλληλων υποστηριγμάτων. Επίσης, μπορούν να εξοπλιστούν με μεγάλους φακούς όταν είναι αναγκαία παρακολούθηση απομακρυσμένων σημείων. Στη συνέχεια έχουμε τις bullet κάμερες, οι οποίες πήραν το όνομά τους από το κυλινδρικό τους σχήμα. Είναι λιγότερο ογκώδεις από τις box κάμερες, αλλά λόγω του μεγέθους τους έχουν περιορισμούς στους φακούς που μπορούν να λάβουν και γι αυτό είναι κατάλληλες για παρακολούθησεις μικρών και μεσαίων αποστάσεων με την ευκρίνεια της εικόνας που παράγουν να είναι συνήθως υποδεέστερη από αυτήν των box καμερών. Στην τελευταία κατηγορία συναντάμε τις κρυφές κάμερες οι οποίες είναι κατασκευασμένες με τέτοιο τρόπο, έτσι ώστε να μην προδίδουν ότι είναι κάμερες παρακολούθησης. Λόγο του μικρού τους μεγέθους οι κάμερες αυτές έχουν χαμηλή ευκρίνεια και είναι αποτελεσματικές μόνο σε παρακολούθησεις μικρών αποστάσεων.



Εικόνα 2: Κάμερες Ασφαλείας

Σε μεγάλο βαθμό οι κάμερες ασφαλείας λειτουργούν σε συνθήκες χωρίς φωτισμό. Αυτό καθίσταται δυνατό μέσω των εξής τεχνολογιών: α. Image intensification β. active illumination γ. thermal, τις οποίες θα αναλύσουμε εκτενέστερα στο επόμενο κεφάλαιο.

3.2 Κάμερες και νυχτερινή λήψη

Η νυχτερινή λήψη είναι η τεχνολογία που μας παρέχει το θαύμα της παρακολούθησης σε απόλυτο σκοτάδι και της βελτίωσης της εικόνας σε περιβάλλον με χαμηλό φωτισμό. Η νυχτερινή λήψη μπορεί να χωριστεί σε τρεις μεγάλες κατηγορίες. Αρχικά έχουμε την image intensification τεχνολογία η οποία λειτουργεί με βάση την αρχή της μεγιστοποίησης της ποσότητας των ληφθέντων φωτονίων από διάφορες φυσικές πηγές όπως το φως των αστεριών και του φεγγαριού. Παραδείγματα αυτής της τεχνολογίας είναι τα γυαλιά νυχτερινής όρασης και διάφορα είδη από low light κάμερες. Η δεύτερη κατηγορία είναι η active illumination (ενεργού φωτισμού) η οποία λειτουργεί με βάση την αρχή της εντατικοποίησης της ζεύξης, στην τεχνολογία απεικόνισης με μια ενεργή πηγή φωτισμού στο φάσμα της εγγύς υπέρυθρης ή της βραχέως υπέρυθρης ακτινοβολίας (NIR και SWIR αντίστοιχα). Παραδείγματα αυτής της τεχνολογίας περιλαμβάνουν τις low light κάμερες. Και τέλος έχουμε την thermal απεικόνιση. Σε αυτή την κατηγορία οι κάμερες λειτουργούν ανιχνεύοντας την διαφορά θερμοκρασίας των αντικειμένων που βρίσκονται στο παρασκήνιο σε σχέση με αυτά που βρίσκονται στο προσκήνιο. Στο επόμενο κεφάλαιο θα μελετηθούν τα πλεονεκτήματα και τα μειονεκτήματα της κάθε τεχνολογίας και θα γίνει μια μικρή αναφορά στην κάθε μια.

3.3 Πλεονεκτήματα-Μειονεκτήματα

Θα ξεκινήσουμε συζητώντας την low light απεικόνιση αφού στις μέρες μας, η πιο γνωστή μέθοδος για να πραγματοποιήσουμε νυχτερινή λήψη είναι με τη χρήση ενισχυτών εικόνας.



Εικόνα 3: Παράδειγμα low light απεικόνισης

Αυτή η μέθοδος της νυχτερινής όρασης ενισχύει το διαθέσιμο φως για να επιτευχθεί καλύτερη όραση εστιάζοντας το διαθέσιμο φως για να αυξήσει την ποιότητα της εικόνας. Χρησιμοποιούμε πράσινο χρώμα σε αυτές τις εφαρμογές επειδή το ανθρώπινο μάτι μπορεί να διαφοροποιήσει περισσότερες σκιές πράσινου από κάθε άλλο χρώμα, επιτρέποντας μεγαλύτερη διαφοροποίηση των αντικειμένων στο πλάνο μας.

Πλεονεκτήματα:

1. Εξαιρετική ευαισθησία σε λήψεις με χαμηλό φωτισμό.
2. Υψηλή ανάλυση.
3. Χαμηλό κόστος.
4. Δυνατότητα εντοπισμού ανθρώπων.
5. Ενισχυμένη οπτική απεικόνιση με τη καλύτερη δυνατή επίδοση σε αναγνώριση και ταυτοποίηση.

Μειονεκτήματα:

1. Επειδή βασίζονται σε μεθόδους ενίσχυσης, απαιτείται κάποια ποσότητα από φως άρα η μέθοδος δεν είναι χρήσιμη σε χώρους με απόλυτο σκοτάδι.
2. Έχουν κατώτερη επίδοση σε φωτεινές λήψεις από κάμερες που χρησιμοποιούνται αποκλειστικά σε φωτεινές λήψεις.
3. Έχουν μεγάλη πιθανότητα βλάβης σε περίπτωση που υπάρξει φωτεινό αντικείμενο σε συνθήκες χαμηλού φωτισμού.

Στη συνέχεια θα περάσουμε στις θερμικές κάμερες οι οποίες διαφέρουν από τις low-light στο ότι δεν χρειάζονται καθόλου φως για να λειτουργήσουν.



Εικόνα 4: Παράδειγμα θερμικής κάμερας

Οι θερμικές κάμερες λειτουργούν με βάση την αρχή ότι όλα τα αντικείμενα εκπέμπουν υπέρυθη ακτινοβολία ως συνάρτηση της θερμότητάς τους. Γενικότερα, όσο πιο ζεστό είναι ένα αντικείμενο τόσο μεγαλύτερη ακτινοβολία εκπέμπει.

Πλεονεκτήματα:

1. Υψηλή αντίθεση σε σενάρια χωρίς φως.
2. Ανιχνεύει πολύ εύκολα ανθρώπους και οχήματα.
3. Δεν επηρεάζεται από φωτεινές πηγές.

Μειονεκτήματα:

1. Είναι πολύ ακριβό στην αγορά και στο χειρισμό.
2. Δεν έχει καλή απόδοση σε περιπτώσεις που το αντικείμενο κινείται με μεγάλη ταχύτητα.

Και τέλος οι near infrared απεικόνιση η οποία είναι μια δημοφιλής και αρκετές φορές οικονομική μέθοδος για να πραγματοποιήσουμε νυχτερινή λήψη.



Εικόνα 5: Παράδειγμα κάμερας «Near Infrared Illumination»

Πτυχιακή Εργασία Τμήματος Μηχανικών Πληροφορικής

Σε αυτή τη μέθοδο, χρησιμοποιείται μια συσκευή η οποία είναι ευαίσθητη στην αόρατη υπέρυθη ακτινοβολία, σε συνδυασμό με μια διάταξη φωτισμού υπέρυθρων. Με αυτό τον τρόπο, αυξάνουμε την ποιότητα εικόνας σε σκηνές με χαμηλό φωτισμό, παράγοντας εικόνες ικανοποιητικής ποιότητας.

Πλεονεκτήματα:

1. Έχει το χαμηλότερο κόστος σε σύγκριση με τις άλλες μεθόδους νυχτερινής λήψης.
2. Εξαλείφει τις σκιές και εντοπίζει ανθρώπους και αντικείμενα.
3. Μπορεί να πραγματοποιήσει λήψεις όταν το αντικείμενο κινείται με μεγάλη ταχύτητα.
4. Είναι ικανό να διαπεράσει έντονη βροχή, ομίχλη, καπνούς.

Μειονεκτήματα:

1. Οι κάμερες μπορούν να ανιχνευτούν εύκολα από τρίτους με συσκευές εντοπισμού υπέρυθρων.

4 Τεχνικές Ανίχνευσης Ανθρώπου

Η αναγνώριση της παρουσίας ανθρώπου είναι μια από τις πιο σημαντικές λειτουργίες που μπορούν να εξαχθούν από βίντεο. Παρόλα αυτά, η εργασία αυτή γίνεται πιο περίπλοκη με την παρουσία διαφόρων μεταβλητών στην εικόνα, όπως η φωτεινότητα, ο φωτισμός, η αντίθεση και το φόντο. Σκοπός του κεφαλαίου είναι η συζήτηση για την ανάπτυξη μιας μεθόδου, η οποία θα ανιχνεύει και θα προσδιορίζει την παρουσία ανθρώπου σε μια ακολουθία βίντεο, και θα την διαφοροποιεί από μη-ανθρώπινες κινήσεις. Από την έρευνα που έγινε καταλήξαμε στο συμπέρασμα ότι ο πιο κοινός τρόπος για να επιτευχθεί ανίχνευση ανθρώπου είναι μέσω της ανίχνευσης προσώπου. Το πρόσωπο του ανθρώπου είναι το πιο ξεχωριστό σημείο στο ανθρώπινο σώμα, και αν ανιχνευθεί με ακρίβεια οδηγεί σε επιτυχή ανίχνευση ανθρώπινης σιλουέτας.

4.1 Ανίχνευση Προσώπου



Εικόνα 6: Εικόνα που παρουσιάζει την ανίχνευση προσώπου σε εξωτερικό χώρο

Από τη δεκαετία του 1970 έως και σήμερα έχουν γίνει αρκετές μελέτες στον τομέα της ανίχνευσης προσώπου, και έχουν δημοσιευθεί αρκετές έρευνες όσον αφορά τους αλγόριθμους που χρησιμοποιούνται σε αυτό τον τομέα. Γενικότερα οι αλγόριθμοι ανίχνευσης προσώπου μπορούν να κατηγοριοποιηθούν σε δυο κύρια είδη. Στην πρώτη κατηγορία ανήκουν οι προσεγγίσεις με βάση τα χαρακτηριστικά(feature-based) παραδείγματος χάριν η «bottom-up» προσέγγιση. Στην αυτή την κατηγορία, εξάγονται τα χαρακτηριστικά του προσώπου από μία εικόνα και χειρίζομαστε τις παραμέτρους της, όπως γωνίες, μέγεθος και αποστάσεις. Στη δεύτερη κατηγορία ανήκουν οι προσεγγίσεις με βάση την εικόνα(image-based), για παράδειγμα προσεγγίσεις με βάση την εμφάνιση(appearance-based) αλλά και η προσέγγιση «template matching». Αυτό το είδος αλγορίθμου, βασίζεται σε τεχνικές εκπαίδευσης μηχανής για να αναγνωρίζει συγκεκριμένα σημεία ενδιαφέροντος. Μια μικρή περιγραφή των πιο κοινών προσεγγίσεων ακολουθεί στις επόμενες ενότητες.

4.1.1 Information based (Top-Down)

Σε αυτή τη μέθοδο καταγράφονται οι σχέσεις μεταξύ των χαρακτηριστικών ενός προσώπου για να γίνει η αναπαράσταση του, και κωδικοποιούνται σαν ένα σύνολο κανόνων. Παραδείγματος χάριν, οι άξιες έντασης για το κεντρικό μέρος του προσώπου είναι ομοιόμορφες από τον έναν άνθρωπο στον άλλο. Συνήθως, ένα πρόσωπο εμφανίζεται με συμμετρικά μεταξύ τους μάτια, μια μύτη και στόμα. Επιπροσθέτως, σε μεγάλο αριθμό αλγορίθμων αυτής της κατηγορίας χρησιμοποιείται η κλίμακα «coarse-to-fine». Σύμφωνα με την «coarse-to-fine» ξεκάνει μια αναζήτηση από την μεγαλύτερη κλίμακα (coarsest) και συνεχίζει στις επόμενες μέχρι ότου να εντοπιστεί η βέλτιστη.

4.1.2 Feature Invariant (Bottom-up)

Στην προσέγγιση αυτή, τα χαρακτηριστικά τα οποία μένουν αναλλοίωτα σε ένα πρόσωπο χρησιμοποιούνται για να βοηθήσουν στην ανίχνευσή του. Αυτά τα χαρακτηριστικά πρέπει να μην αλλάζουν σε ποικίλες οπτικές γωνίες, φωτισμούς και διαφορετικές συνθήκες. Οι συνήθεις αλγόριθμοι που ανήκουν σε αυτή την προσέγγιση είναι:

A. Προσέγγιση με Βάση το Χρώμα ή Μοντέλο Δέρματος

Αυτή η μέθοδος αξιοποιεί το γεγονός ότι το χρώμα του δέρματος μπορεί να χρησιμοποιηθεί σαν στοιχείο παρουσίας ανθρώπου, δεδομένου του ότι μια μεγάλη ποικιλία αποχρώσεων του δέρματος από διαφορετικές φυλές μπορεί να καταταχτεί σε μια ενιαία περιοχή. Ο Vezhnevets V. παρουσίασε τέσσερις τεχνικές μοντελοποίησης με βάση τα pixel: την μη παραμετρική μοντελοποίηση διανομής του δέρματος (non-parametric skin distribution modelling), την παραμετρική (parametric skin distribution modelling), την δυναμική (dynamic skin distribution modelling) και τέλος την ρητώς οριζόμενη περιοχή του δέρματος (explicitly defined skin region).

B. Προσέγγιση με βάση τα Χαρακτηριστικά Προσώπου

Αυτή η μέθοδος στην οποία, χρησιμοποιούνται τα παγκόσμια χαρακτηριστικά του ανθρώπινου σώματος, π.χ. δέρμα, μέγεθος, και σχήμα, αλλά και τα λεπτομερή, π.χ. μάτια, μύτη, χείλη, έγινε αρκετά δημοφιλής τα τελευταία χρόνια. Στις περισσότερες περιπτώσεις προκειμένου να επιτευχθεί η ανίχνευση προσώπου χρησιμοποιούνται τα παγκόσμια χαρακτηριστικά για να εντοπιστεί η υποψήφια περιοχή και στη συνέχεια γίνεται επαλήθευση με την χρήση των λεπτομερών χαρακτηριστικών.

Γ. Υφής (Texture)

Το ανθρώπινο πρόσωπο διαφέρει από τα άλλα αντικείμενα σε υφή. Σε αυτή τη μέθοδο, εξετάζουμε την πιθανότητα η υποψήφια περιοχή να ανήκει σε υφή ανθρώπινου προσώπου, με τη χρήση ενός πίνακα Space Gray Level Dependence (SGLD matrix).

4.1.3 Template matching methods

Αυτές οι μέθοδοι βασίζονται στη μέτρηση του βαθμού ομοιότητας μεταξύ της υποψηφίας περιοχής και του αποθηκευμένου μοτίβου του προσώπου. Η προκαθορισμένη εικόνα μπορεί να είναι για όλο το μοτίβο του προσώπου ή για μοναδικά χαρακτηριστικά του, όπως είναι τα μάτια, τα χείλη και η μύτη. Οι πιο συνήθεις αλγόριθμοι σε αυτή την προσέγγιση είναι:

A. Προκαθορισμένα Πρότυπα Προσώπου

Σε αυτή την κατηγορία αλγόριθμοι πολλά πρότυπα από ολόκληρο το πρόσωπο ή κομμάτια αυτού αποθηκεύονται με σκοπό να αναλυθούν τα μοτίβα του, για μετέπειτα χρήση τους.

B. Παραμορφώσιμα Πρότυπα

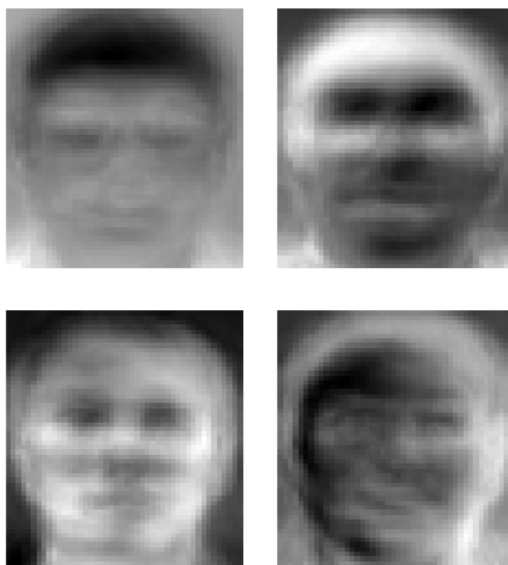
Σε αυτή την κατηγορία ένα ελαστικό μοντέλο των χαρακτηριστικών του προσώπου χρησιμοποιείται ως μοντέλο αναφοράς, στο οποίο το παραμορφώσιμο πρότυπο του αντικειμένου ενδιαφέροντος ταιριάζεται.

4.1.4 Appearance based

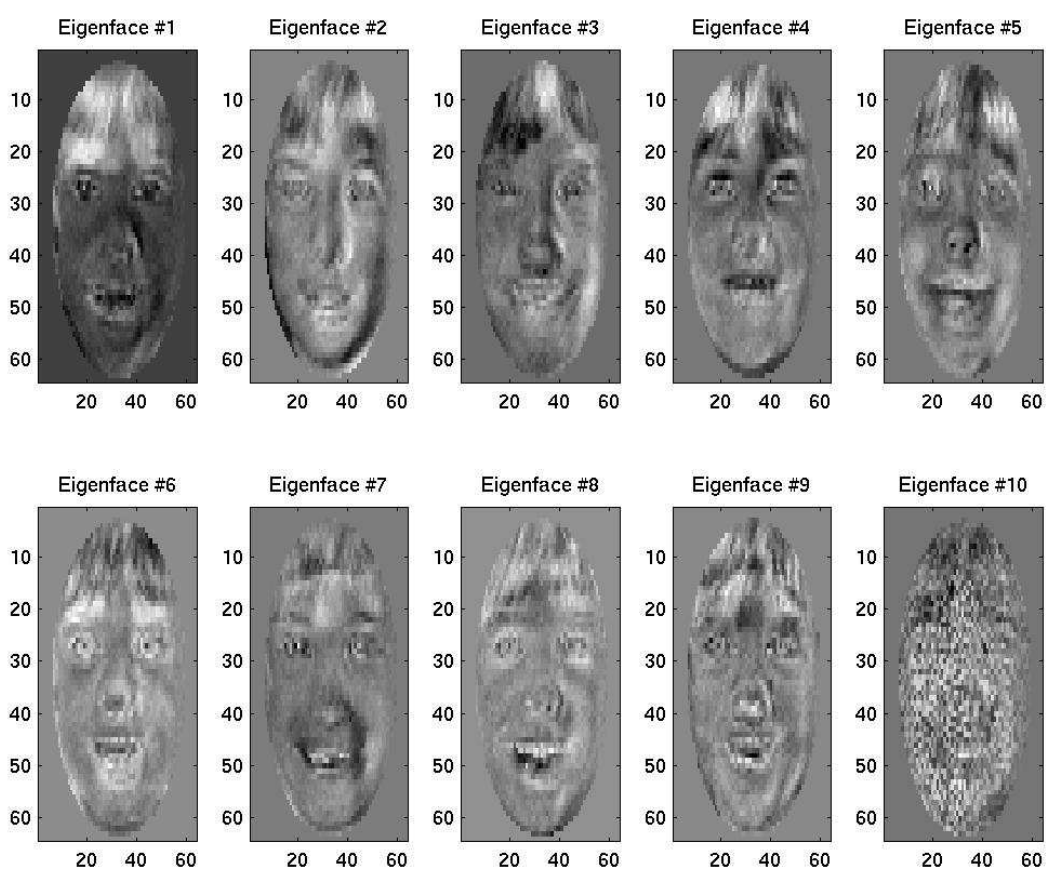
Σε αντίθεση με τις «template matching» μεθόδους στις οποίες τα πρότυπα έχουν προκαθοριστεί από ειδικούς, οι appearance-based μέθοδοι, μαθαίνουν τα πρότυπα από ένα σετ εικόνων, χρησιμοποιώντας στατιστική ανάλυση και εκμάθηση μηχανής. Μερικά παραδείγματα αλγόριθμων που χρησιμοποιούν αυτή τη μέθοδο είναι:

A. Eigenfaces

Ο συγκεκριμένος αλγόριθμος είναι ένα σετ από ιδιοδιανύσματα που χρησιμοποιούνται στην τεχνητή όραση για να αντιμετωπίσουν το πρόβλημα της ανίχνευσης προσώπου. Αυτά τα ιδιοδιανύσματα προκύπτουν από τον πίνακα διακύμανσης της κατανομής των πιθανοτήτων για τα εκάστου ανθρώπινα πρόσωπα. Ουσιαστικά ένα σύνολο από eigenfaces μπορεί να παραχθεί από την εκτέλεση μιας διαδικασίας, που ονομάζεται μαθηματική ανάλυση κύριων συνιστωσών σε ένα μεγάλο σύνολο εικόνων που απεικονίζουν διαφορετικά ανθρώπινα πρόσωπα. Κάθε ανθρώπινο πρόσωπο μπορεί να θεωρηθεί ότι είναι ένας συνδυασμός από τυποποιημένα πρόσωπα. Αξίζει να σημειωθεί ότι δεν χρειαζόμαστε πολλά eigenfaces για να επιτύχουμε μια σωστή προσέγγιση των περισσότερων προσώπων, επίσης, επειδή τα πρόσωπα αποθηκεύονται σαν μια απλή λίστα τιμών χρειαζόμαστε πολύ λιγότερο χώρο σε μια βάση δεδομένων για την αποθήκευσή τους.



Εικόνα 7: Παράδειγμα 1 από Eigenfaces



Εικόνα 8: Παράδειγμα 2 από Eigenfaces

B. Distributed-based

Εδώ το καταναμημένο μοτίβο ενός αντικειμένου μαθαίνεται με τη χρήση ενός σετ θετικών και αρνητικών εικόνων του αντικειμένου.

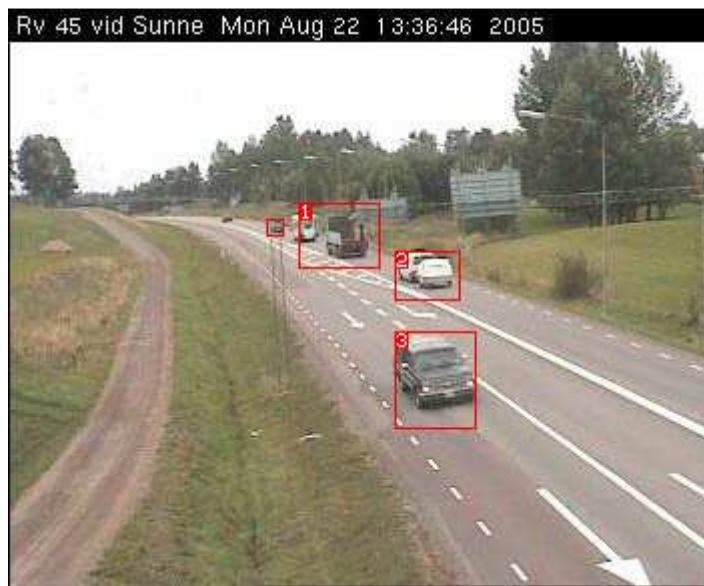
Γ. Neural Network

Τα νευρωνικά δίκτυα(κόμβοι) λειτουργούν για να παράγουν παράλληλες συναρτήσεις. Η ιδέα των νευρωνικών δικτύων ήρθε από το κεντρικό νευρικό σύστημα. Παρόλα αυτά, τα δίκτυα αυτά εκπαιδεύτηκαν για να ανιχνεύουν την παρουσία προσώπου αφού λαμβάνουν δείγματα από πρόσωπο.

Δ. Support Vector Machine

Τα «Support Vector Machines» είναι εκπαιδευόμενες μηχανές οι οποίες κάνουν δυαδικούς συσχετισμούς. Η ιδέα είναι να μεγιστοποιηθεί το περιθώριο μεταξύ θετικών και αρνητικών σετ διανυσμάτων προκειμένου να επιτευχθεί ένα βέλτιστο όριο το οποίο θα διαχωρίζει τα δυο σετ διανυσμάτων.

4.2 Ανίχνευση Κίνησης



Εικόνα 9: Παράδειγμα αλγορίθμου ανίχνευσης κίνησης

Σε αντίθεση με τις ακίνητες εικόνες, οι ακολουθίες βίντεο έχουν περισσότερες πληροφορίες σχετικά με το ιστορικό των κινούμενων αντικειμένων, το οποίο βοηθά στο διαχωρισμό του φόντου από το προσκήνιο. Η ανίχνευση κινούμενων αντικειμένων είναι πολύ σημαντική σε βίντεο παρακολούθησης. Σε πολλά περιβάλλοντα η κίνηση μπορεί να είναι είτε ενδιαφέρουσα, π.χ. άνθρωπος (salient- εμφανής), είτε μη ενδιαφέρουσα, π.χ. ταλαντευόμενα κλαδιά. Η ανίχνευση κίνησης σε βίντεο θεωρείται ένα σημαντικό και αρκετά δύσκολο πρόβλημα. Σε πολλά πραγματικά περιβάλλοντα η κίνηση προκαλείται και από ενδιαφέροντα αλλά και από μη ενδιαφέροντα ερεθίσματα. Σαν εμφανή κίνηση θεωρούμε την κίνηση που προκύπτει από ένα τυπικό αντικείμενο παρακολούθησης, π.χ. άνθρωπο ή όχημα, σε αντίθεση με τις άλλες παραπλανητικές κινήσεις, όπως είναι οι ακτίνες του φωτός όταν πέφτουν στο νερό ή το ανέμισμα των φύλων ενός δέντρου. Οι παραπλανητικές αυτές

κινήσεις, σε ένα πραγματικό περιβάλλον, κάνουν το πρόβλημα της ανίχνευσης κίνησης ακόμη πιο δύσκολο.

Γενικότερα, οι κινούμενες περιοχές ανιχνεύονται με τον εντοπισμό των αλλαγών που δημιουργούνται στην ακολουθία εικόνων. Οι περισσότερες εφαρμογές που έχουν δημιουργηθεί πάνω στην ανίχνευση κίνησης, εφαρμόζουν ορισμένα στάδια προεπεξεργασίας πριν εφαρμόσουν τον αλγόριθμο ανίχνευσης αλλαγών. Αυτά τα στάδια περιέχουν γεωμετρικές προσαρμογές αλλά και προσαρμογές έντασης. Το πρόβλημα της ποικιλίας στην ένταση του φωτός, λύνεται με την προσαρμογή της. Σε αυτήν την προσαρμογή το αποτέλεσμα του φωτισμού μειώνεται ανάλογα με τη μέθοδο που θα χρησιμοποιήσουμε. Ο Elgammal σε έρευνα του δηλώνει ότι αν μετασχηματίσουμε τις τιμές RGB σε «chromatic colour space», το μοντέλο μας μένει ανεπηρέαστο από μικρές αλλαγές στο φως.

Υπάρχουν αρκετοί τρόποι για να εντοπίσεις μια αλλαγή σε ένα βίντεο. Πρόσφατες μελέτες έδειξαν ότι η μέθοδος διαφοροποίησης εικόνας (image differencing method) είναι πιο αποτελεσματική από τις υπόλοιπες στην ανίχνευση αλλαγών. Η αφαίρεση του φόντου είναι μια συμβατική και αποτελεσματική προσέγγιση για να ανιχνεύσεις κινούμενα αντικείμενα όταν το φόντο είναι ακίνητο. Για να ανιχνεύσεις όμως κινούμενα αντικείμενα σε μια δυναμική σκηνή, έχουν αναπτυχθεί προσαρμοστικές τεχνικές αφαίρεσης φόντου. Ο Ren πρότεινε μια χωρική κατανομή από «Gaussian» μοντέλα, για να αντιμετωπίσει την ανίχνευση κινούμενων αντικειμένων, με αντιστάθμιση κίνησης, όπου υπάρχει. Τα αποτελέσματά του επέδειξαν την ικανότητα του συστήματος να ανιχνεύει πολύ μικρά αντικείμενα, τα οποία κινούνταν σε ένα φόντο με έντονη υφή. Ο Stuffer μοντελοποίησε κάθε pixel σαν μια μίξη από «Gaussians» και χρησιμοποίησε μια διαδικτυακή προσέγγιση για να ανανεώνει το μοντέλο. Το σύστημά του μπορούσε να αντιμετωπίσει αλλαγές στο φωτισμό, αντικείμενα που κινούνται με μικρή ταχύτητα και οποιαδήποτε αλλαγή στη σκηνή, που προέκυπτε από την πρόσθεση ή την αφαίρεση αντικειμένων. Ο Monpet πρότεινε μια διαδικτυακή μέθοδο με βάση την πρόβλεψη κίνησης, για την μοντελοποίηση δυναμικών σκηνών. Η μέθοδός του δοκιμάστηκε σε μια ακτή με κύματα και κινούμενα φύλλα δέντρων. Τα αποτελέσματα δεν ήταν ενθαρρυντικά αφού χρειάζονταν εκατοντάδες εικόνες της σκηνής χωρίς κινούμενα αντικείμενα, προκειμένου το σύστημα να καταφέρει να μάθει το φόντο. Επιπλέον, τα κινούμενα αντικείμενα δεν μπορούσαν να ανιχνευθούν αν κινούνταν στην ίδια κατεύθυνση με αυτή των κυμάτων της θάλασσας. Πρόσφατα, κάποιιοι υβριδικοί ανιχνευτές αναπτύχθηκαν, οι οποίοι συνδυάζουν μια απεικόνιση χρονικής διαφοράς με ένα προσαρμοστικό μοντέλο αφαίρεσης φόντου, για να αντιμετωπίσουν τις αλλαγές του φωτός.

Παρόλα αυτά, καμία από τις υπάρχουσες τεχνικές δεν μπορεί να προσαρμοστεί σε γρήγορες αλλαγές στην εικόνα, όπως είναι η αλλαγή ενός φωτός που αναβοσβήνει. Τα σημαντικότερα μειονεκτήματα των τεχνικών αυτών συνοψίζονται ως εξής:

- Χρειάζονται εκατοντάδες εικόνες για να μάθουν το φόντο.
- Δεν μπορούν να χειριστούν γρήγορες αλλαγές και παραπλανητικές κινήσεις.
- Δεν αφήνουν περιθώρια σε ακίνητα αντικείμενα να αρχίσουν να μετακινούνται.

Ποικίλες μέθοδοι για την ανίχνευση εμφανούς κίνησης έχουν αναπτυχθεί. Ο Wildes πρότεινε μια μέτρηση της ενδιαφέρουσας κίνησης με τη χρήση χωροχρονικού φιλτραρίσματος. Η μέτρηση αυτή υποθέτει ότι το αντικείμενο κινείται με συγκεκριμένη ταχύτητα, εξαιτίας της εξαρτώμενης από την ταχύτητα φύσεως των χωροχρονικών φίλτρων. Αυτή η μέθοδος δεν λειτούργησε για αντικείμενα τα οποία κινούνται με πολύ μικρή ταχύτητα. Ο Wixson παρουσίασε μια μέθοδο για ανίχνευση κίνησης συσσωρεύοντας την ευθυγραμμισμένη ροή. Υπολόγισαν την οπτική ροή των υπό-pixel και ενσωμάτωσαν την οπτική ροή από frame σε frame του κάθε pixel, για να υπολογίσουν μια ακριβή προσέγγιση του πόσο έχει μετακινηθεί ολόκληρη η εικόνα. Σε κάθε frame ανανέωναν την τιμή η οποία ήταν άμεσα συνδεδεμένη με την απόσταση που έχει ταξιδέψει ένα σημείο. Η μέθοδός τους

όμως ήταν πολύ χρονοβόρα και άφηνε ίχνη από τα κινούμενα αντικείμενα με το πέρασμα του χρόνου.

4.3 Ανίχνευση Δέρματος



Εικόνα 10: Ανίχνευση Δέρματος

Το χρώμα και η υφή του δέρματος αποτελούν σημαντικές ενδείξεις που οι άνθρωποι χρησιμοποιούν συνειδητά ή ασυνείδητα για να συμπεράνουν μια ποικιλία από πολιτισμικές πτυχές των υπολοίπων. Το χρώμα του δέρματος και η υφή μπορεί να είναι μια ένδειξη της φυλής, της ηλικίας, και της ομορφιάς του καθενός από εμάς, αλλά οι ερμηνείες ποικίλουν ανά τους διαφορετικούς πολιτισμούς και τις χρονικές περιόδους. Στις εικόνες και στα βίντεο όμως, το χρώμα του δέρματος είναι ένα σημάδι παρουσίας ανθρώπου. Γι αυτό το λόγο, τις τελευταίες δυο δεκαετίες αναλυτική έρευνα έχει εστιάσει στην ανίχνευση δέρματος. Η ανίχνευση δέρματος είναι η διαδικασία με την οποία εντοπίζουμε pixel και περιοχές που έχουν το χρώμα του δέρματος. Αυτή η διαδικασία χρησιμοποιείται συνήθως ως ένα βήμα προεπεξεργασίας για να βρούμε τις περιοχές που περιέχουν ενδεχομένως ανθρώπινα πρόσωπα και άκρα σε εικόνες. Διαφορετικές προσεγγίσεις τεχνητής όρασης έχουν αναπτυχθεί για την ανίχνευση δέρματος. Ένας ανιχνευτής δέρματος συνήθως μετασχηματίζει ένα δεδομένο pixel σε ένα κατάλληλο χώρο χρωμάτων(color space) και στη συνέχεια χρησιμοποιεί ένα ταξινομητή δέρματος για να τα ονομάσει σε pixel δέρματος ή μη-δέρματος. Ένας ταξινομητής καθορίζει ένα πεδίο στο χώρο χρωμάτων ως χρώματα δέρματος βασισμένοι σε μια εκπαιδευόμενη βάση δεδομένων.

Όπως θα αναλύσουμε σε αυτό το κεφάλαιο, η ανίχνευση δέρματος είναι ένα υπολογιστικά εύκολο έργο και μπορεί να πραγματοποιηθεί πολύ αποτελεσματικά, ένα χαρακτηριστικό που ενθαρρύνει τη χρήση της ανίχνευσης δέρματος σε πολλές εφαρμογές που ασχολούνται με ανάλυση βίντεο. Παρόλα αυτά, η ανίχνευση pixel στον τόνο του δέρματος, αποδείχτηκε ότι είναι αρκετά περίπλοκη για πολλούς άλλους λόγους. Η εμφάνιση του δέρματος σε μια εικόνα εξαρτάται από τις συνθήκες του φωτισμού του μέρους που έγινε η λήψη. Εμείς σαν άνθρωποι, είμαστε πολύ καλοί στο να αναγνωρίζουμε τα χρώματα των αντικειμένων σε οποιαδήποτε φωτισμό. Ως εκ τούτου, μια σημαντική πρόκληση στην

ανίχνευση δέρματος είναι να αναπαραστήσουμε το χρώμα με ένα τρόπο με τον οποίο να μένει αναλλοίωτο ή τουλάχιστον να μην είναι πολύ ευαίσθητο στις αλλαγές του φωτισμού. Η επιλογή του χώρου χρωμάτων επηρεάζει πάρα πολύ την απόδοση οποιουδήποτε ανιχνευτή δέρματος, και την ευαισθησία του στις αλλαγές του φωτισμού. Άλλη μια πρόκληση αποτελούν τα αντικείμενα που έχουν τον ίδιο τόνο με το ανθρώπινο δέρμα, όπως είναι το ξύλο, το δέρμα, η άμμος κτλ. Αυτό προκαλεί, πολλές λάθος ανιχνεύσεις, σε οποιονδήποτε ανιχνευτή δέρματος, σε μη ελεγχόμενα περιβάλλοντα.

Η διαδικασία της ανίχνευσης δέρματος έχει δυο φάσεις. Τη φάση της εκμάθησης, και τη φάση της ανίχνευσης. Το στάδιο της εκπαίδευσης περιλαμβάνει τρία βασικά στάδια:

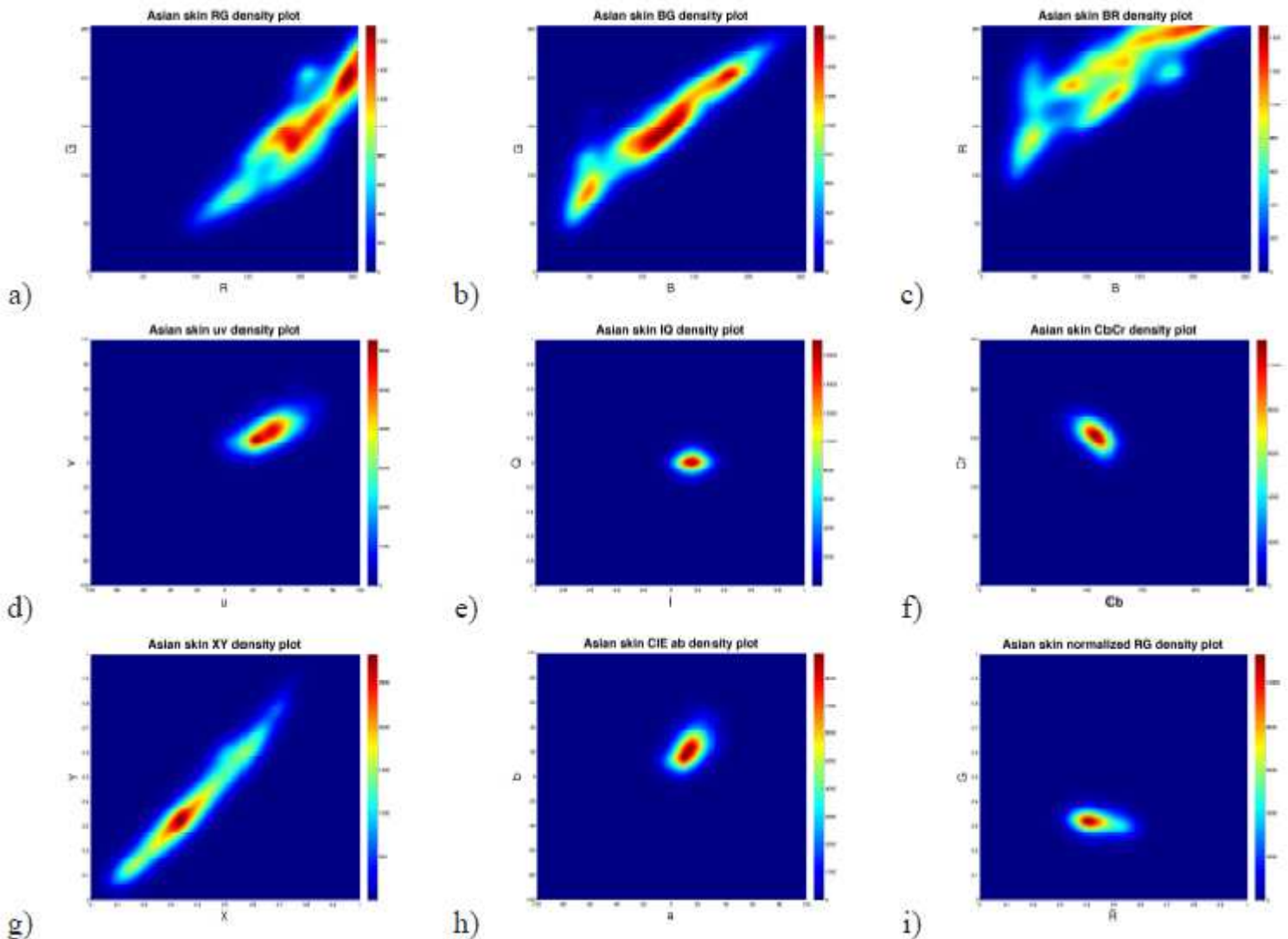
1. Συλλογή μιας βάσης δεδομένων με τόνους δέρματος από διαφορετικές εικόνες. Μια τέτοια βάση συνήθως περιέχει τόνους από μια μεγάλη ποικιλία ανθρώπων σε διαφορετικές καταστάσεις φωτός.
2. Να διαλέξουμε τον κατάλληλο χώρο χρώματος
3. Εκμάθηση των παραμέτρων ενός ταξινομητή δέρματος

Σε κάθε εκπαιδευόμενο ανιχνευτή δέρματος, ο προσδιορισμός των πιξελ του δέρματος γίνεται σε τρεις φάσεις:

1. Τη μετατροπή της εικόνας στον ίδιο χώρο χρωμάτων που χρησιμοποιήσαμε στο στάδιο της εκμάθησης
2. Να κατατάξουμε τα pixel με τη χρήση του ταξινομητή δέρματος σε δέρμα ή μη-δέρμα.
3. Τέλος, τα pixel επεξεργάζονται μορφολογικά για να επιβληθεί χωρική ομοιογένεια στις περιοχές που εντοπίστηκαν.

Σε οποιοδήποτε χώρο χρωμάτων, το χρώμα του δέρματος καταλαμβάνει ένα συμπαγές και μεγάλο μέρος στο χώρο. Μια τέτοια περιοχή ονομάζεται σύμπλεγμα χρώματος δέρματος. Ένας ταξινομητής δέρματος είναι ένα πρόβλημα ταξινόμησης πρώτης ή δεύτερης κατηγορίας. Ένα pixel κατηγοριοποιείται και επισημαίνεται είτε πρόκειται για δέρμα ή μη δέρμα, δεδομένου ενός μοντέλου δέρματος σε ένα συγκεκριμένο χώρο χρώματος. Στο πλαίσιο της ταξινόμησης του δέρματος τα αληθώς θετικά είναι pixel δέρματος που ο ταξινομητής ορθώς επισύναψε ως δέρμα. Τα αληθώς αρνητικά είναι περιοχές μη δέρματος που ο ταξινομητής ορθώς επισύναψε ως μη δέρμα. Όμως κάθε ταξινομητής κάνει λάθη. Μπορεί να επισυνάψει μια περιοχή που δεν είναι δέρμα σαν δέρμα και το αντίστροφο. Αυτά τα λάθη ονομάζονται ψευδώς θετικά στην πρώτη περίπτωση και ψευδώς αρνητικά στη δεύτερη, αλλά και τα δυο οδηγούν σε λάθος ανίχνευση. Ένας καλός ταξινομητής πρέπει να έχει χαμηλά ποσοστά ψευδώς θετικών και αρνητικών επισυνάψεων.

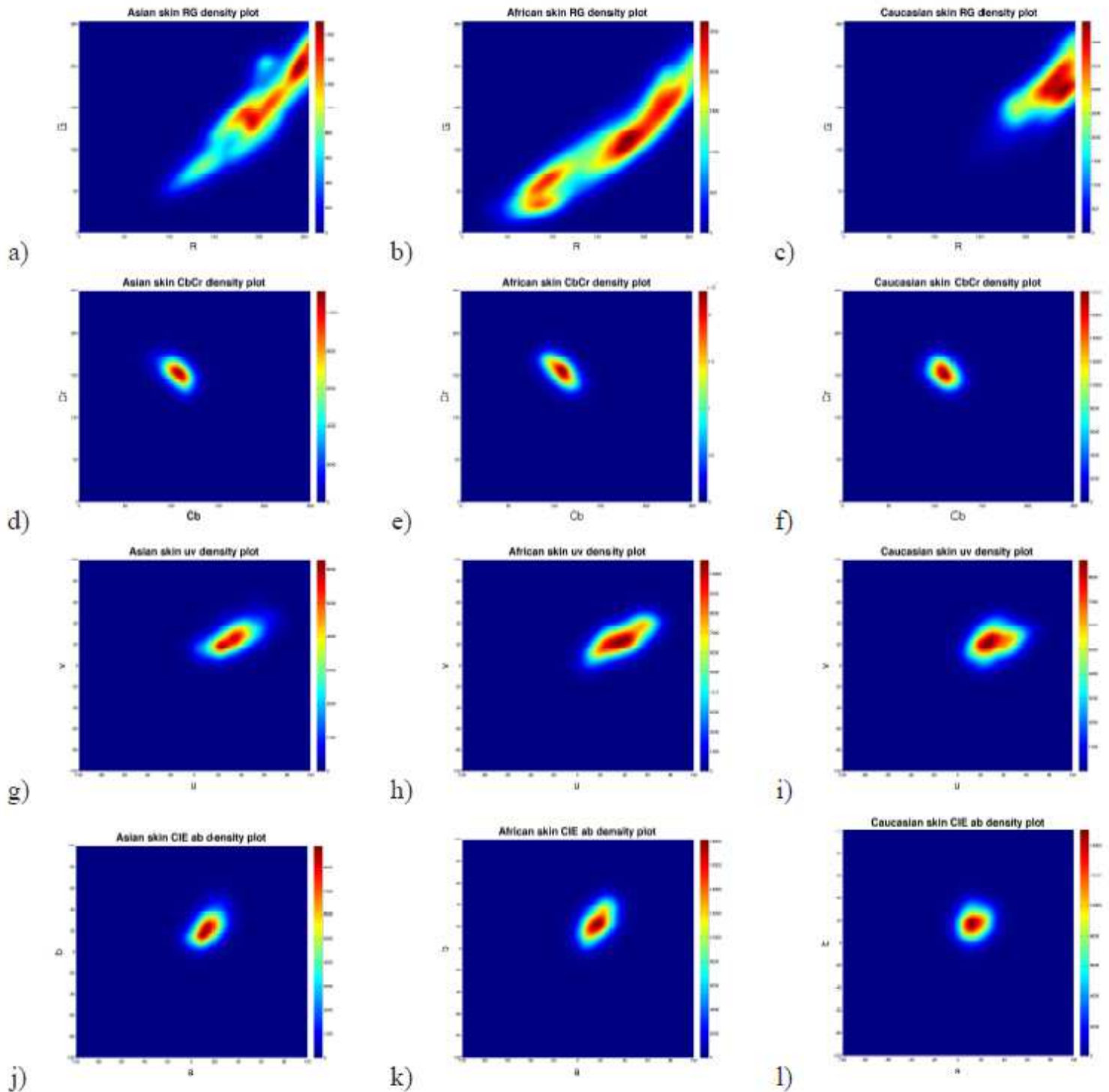
Το ίδιο ισχύει και στην ανίχνευση δέρματος, γι αυτό το λόγο η επιλογή του χώρου χρωμάτων είναι μια πολύ σημαντική πτυχή της ανίχνευσης δέρματος. Το χρώμα χρειάζεται να αναπαρασταθεί σε ένα χώρο χρωμάτων στον οποίο η τάξη του δέρματος είναι πιο συμπαγής έτσι ώστε να μπορεί να ταξινομηθεί πιο εύκολα. Η επιλογή του χώρου χρωμάτων επηρεάζει άμεσα και τον ταξινομητή που θα χρησιμοποιήσουμε.



Εικόνα 11: Διάγραμμα πυκνότητας δέρματος Ασιατών σε διαφορετικούς χρωματικούς χώρους

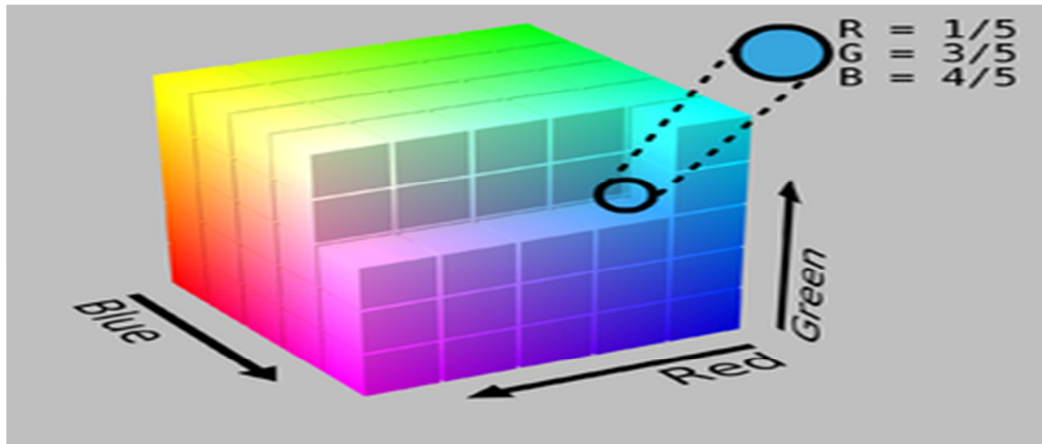
4.3.1 Ανίχνευση δέρματος και χώροι χρωμάτων

Οι Forsyth και Fleck επισήμαναν ότι το χρώμα του ανθρώπινου δέρματος έχει ένα περιορισμένο εύρος αποχρώσεων αφού η εμφάνιση του δέρματος προκύπτει από ένα συνδυασμό του κόκκινου του αίματος και της μελανίνης. Γι αυτό το λόγο, το χρώμα του δέρματος δεν συμπίπτει τυχαία με ένα οποιοδήποτε χώρο χρωμάτων αλλά ομαδοποιείται σε μια μικρή περιοχή του χρωματικού χώρου. Μια μεγάλη ποικιλία από χρωματικούς χώρους χρησιμοποιείται προκειμένου να βρεθεί ο κατάλληλος στον οποίο το χρώμα του δέρματος κάτω από διαφορετικές συνθήκες φωτισμού θα παραμένει ανεπηρέαστο. Σε αυτή την ενότητα θα περιγράψουμε ορισμένους χρωματικούς χώρους που χρησιμοποιούνται πιο συχνά στην ανίχνευση δέρματος και θα συζητήσουμε τον τρόπο με τον οποίο την επηρεάζουν.



Εικόνα 12: Διάγραμμα πυκνότητας δέρματος Ασιατών, Αφρικανών και Καυκάσιων σε διαφορετικούς χρωματικούς χώρους.

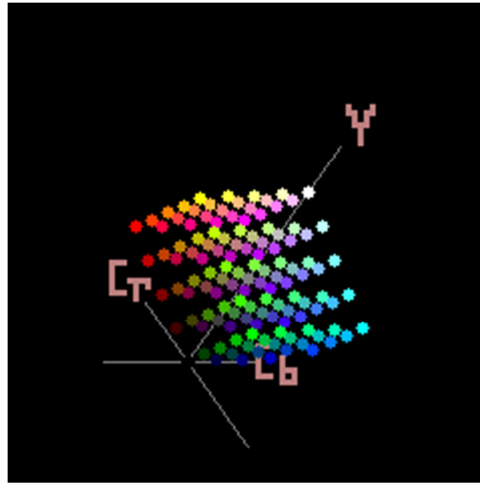
4.3.1.1 RGB Χρωματικός Χώρος



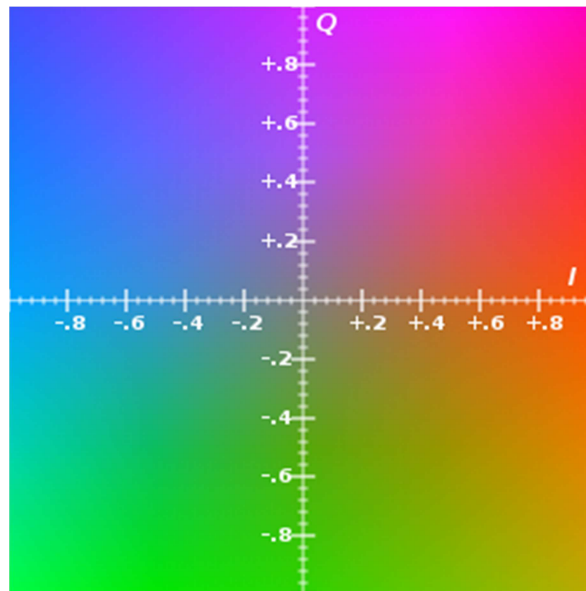
Εικόνα 13: RGB χρωματικός χώρος

Ο RGB είναι ο πιο ευρέως χρησιμοποιούμενος χρωματικός χώρος στις ψηφιακές εικόνες. Κωδικοποιεί τα χρώματα σαν ένα συνδυασμό των τριών βασικών χρωμάτων, του κόκκινου(Red), του πράσινου(Green) και του μπλε(Blue). Ο RGB χρωματικός χώρος συνήθως απεικονίζεται σαν ένας τρισδιάστατος κύβος στον οποίο τα τρία αυτά χρώματα είναι οι 3 άξονες. Ένα κύριο πλεονέκτημα του RGB είναι η απλότητά του. Όμως, δεν είναι ομοιομόρφως αντιληπτά κάτι το οποίο σημαίνει ότι οι αποστάσεις του RGB δεν αντιστοιχούν γραμμικά με την ανθρώπινη αντίληψη. Επιπλέον, ο συγκεκριμένος χρωματικός χώρος δεν διαχωρίζει τη φωτεινότητα και τη χρωματικότητα και οι τρεις συνιστώσες συσχετίζονται σε πολύ μεγάλο βαθμό. Μια οποιαδήποτε απόχρωση στο RGB αποτελείται από ένα συνδυασμό των βασικών αυτών χρωμάτων γι αυτό το λόγο αν αλλάξουμε το φωτισμό σε ένα κομμάτι δέρματος επηρεάζονται και οι τρεις συνιστώσες. Με άλλα λόγια, το μέρος στο οποίο βρίσκουμε το χρώμα του δέρματος στο κύβο RGB αλλάζει ανάλογα με τη φωτεινότητα στην οποία έγινε η λήψη. Αυτό έχει σαν αποτέλεσμα, ένα ευρύ φάσμα από αποχρώσεις του δέρματος στον RGB κύβο. Αυτό γίνεται αντιληπτό στην πρώτη γραμμή της εικόνας 11, στον οποίο μοντέλα δέρματος από εικόνες ασιατών, οι οποίες τραβήχτηκαν κάτω από τυχαίες συνθήκες φωτισμού, σχεδιάστηκαν στον RGB χρωματικό χώρο. Η ομαδοποίηση του χρώματος του δέρματος επεκτείνεται σε ένα μεγάλο μέρος του χρωματικού χώρου προκειμένου να καλύψει τις αποχρώσεις που προκύπτουν από διαφορετικούς φωτισμούς. Ομοίως, οι ομάδες χρωμάτων από διαφορετικές φυλές εντοπίζονται σε διαφορετικά σημεία του RGB κύβου. Αυτό μπορούμε να το δούμε στην πρώτη σειρά της εικόνας 12. Παρ' όλους τους περιορισμούς, το RGB χρησιμοποιείται στο θεωρητικό κομμάτι της ανίχνευσης δέρματος εξαιτίας της απλότητάς του.

4.3.1.2 Χρωματικός χώρος TV



Εικόνα 14: Αναπαράσταση του $YCbCr$ χρωματικού χώρου

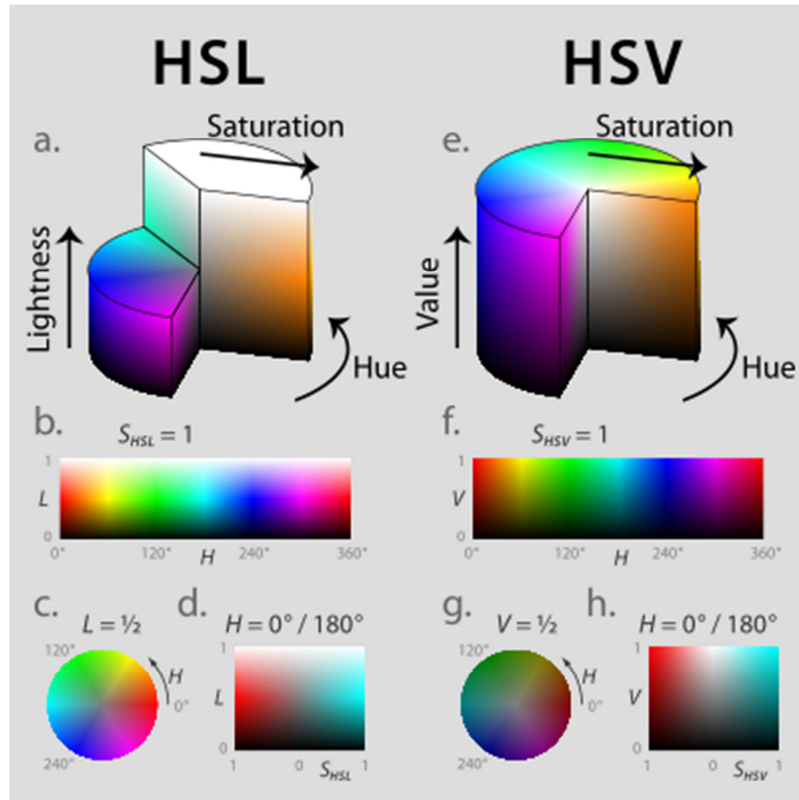


Εικόνα 15: Αναπαράσταση του YIQ χρωματικού χώρου

Μια διαφορετική κατηγορία χώρου χρωμάτων είναι οι ορθογώνιες, οι οποίες χρησιμοποιούνται στην τηλεοπτική μετάδοση. Σε αυτή περιλαμβάνονται οι YUV , YIQ και $YCbCr$. Η YIQ χρησιμοποιείται στην NTSC τηλεοπτική εκπομπή ενώ η $YCbCr$ χρησιμοποιείται στο JPEG πρότυπο συμπίεσης εικόνας και στο MPEG πρότυπο συμπίεσης βίντεο. Ένα πλεονέκτημα που προκύπτει από τη χρήση αυτού του χρωματικού χώρου είναι ότι τα περισσότερα βίντεο είναι ήδη κωδικοποιημένα σε αυτό το χρωματικό χώρο. Ο μετασχηματισμός από RGB σε οποιοδήποτε από αυτούς τους χρωματικούς χώρους είναι μια απλή γραμμική μετατροπή. Όλοι αυτοί οι χρωματικοί χώροι διαχωρίζουν το κανάλι του φωτισμού (Y) από τα δυο ορθογώνια χρωματικά κανάλια (UV , IQ , C_bC_r). Συνεπώς, σε αντίθεση με το RGB, η τοποθεσία που βρίσκεται το χρώμα του δέρματος στο κανάλι χρώματος δεν επηρεάζεται από αλλαγή της έντασης της φωτεινότητας όπως βλέπουμε και στην εικόνα 12-d,e,f. Αυτό μας διευκολύνει και στην κατασκευή ενός ανιχνευτή δέρματος ο οποίος είναι αμετάβλητος από την ένταση της φωτεινότητας και χρησιμοποιεί πιο απλούς ταξινομητές. Η πυκνότητα του χρώματος του δέρματος πάνω στα χρωματικά κανάλια μπορεί εύκολα να προσεγγιστεί με τη χρήση μιας πολυμεταβλητής Gaussian κατανομής. Επιπλέον, τα χρώματα του δέρματος διαφορετικών φυλών σχεδόν συμπίπτουν στα χρωματικά κανάλια,

όπως φαίνεται στην δεύτερη και τρίτη σειρά της εικόνας 12. Ως εκ τούτου, η χρήση ενός τέτοιου χρωματικού χώρου έχει σαν αποτέλεσμα ανιχνευτές δέρματος, οι οποίοι είναι ανεξάρτητοι από τις πιθανές αποχρώσεις του ανθρώπινου δέρματος. Η απλότητα αυτού του μετασχηματισμού καθώς και οι ιδιότητες που περιγράψαμε καθιστούν αυτούς τους χρωματικούς χώρους ευρέως χρησιμοποιούμενους σε εφαρμογές ανίχνευσης δέρματος.

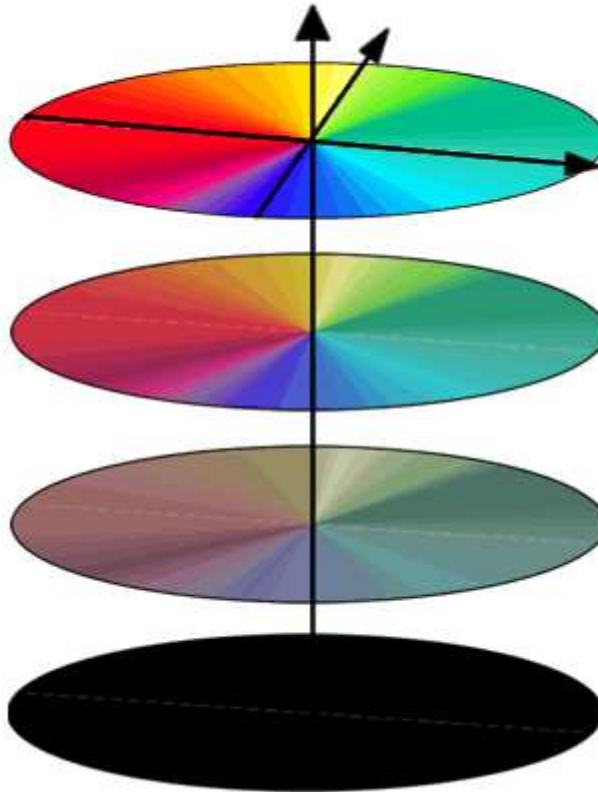
4.3.1.3 Αντιληπτικός χρωματικός χώρος



Εικόνα 16: HSL και HSV.

Αντιληπτικοί χρωματικοί χώροι, όπως είναι οι HIS, HSV/HSB και HSL(HLS) είναι εξίσου δημοφιλείς στην ανίχνευση δέρματος. Αυτοί οι χρωματικοί χώροι διαχωρίζουν τα τρία συστατικά: την απόχρωση (H), τον κορεσμό (S) και τη φωτεινότητα (I,V ή L). Ουσιαστικά, οι HSV χρωματικοί χώροι είναι μετασχηματισμοί του RGB χρωματικού κύβου, και μπορούν να απεικονιστούν από τον RGB χρωματικό χώρο μέσω ενός μη γραμμικού μετασχηματισμού. Ένα από τα πλεονεκτήματα αυτού του χρωματικού χώρου στην ανίχνευση δέρματος είναι ότι επιτρέπουν στους χρήστες να προσδιορίσουν ενστικτωδώς το όριο της τάξης του χρώματος του δέρματος από άποψη κορεσμού και φωτεινότητας. Όσο για την τελευταία συνιστώσα (I,V,L) η οποία μας δίνει τη φωτεινότητα, συνήθως δεν την λαμβάνουμε τόσο υπόψη, προκειμένου να μειώσουμε την εξάρτηση του χρώματος του δέρματος από την φωτεινότητα.

4.3.1.4 Χρωματομετρικός χρωματικός χώρος



Εικόνα 17: Αναπαράσταση του CIE-lab.

Ο διαχωρισμός της χρωματικότητας από την φωτεινότητα επιτυγχάνεται και στους χρωματομετρικούς χρωματικούς χώρους, όπως είναι οι CIE-XYZ, CIE-xy, CIE-lab. Ο CIE-XYZ χρωματικός χώρος είναι ένας από τους πρώτους μαθηματικώς ορισμένους χρωματικούς χώρους. Είναι βασισμένος σε εκτεταμένες μετρήσεις της ανθρώπινης οπτικής αντίληψης και χρησιμεύει ως βάση για πολλούς άλλους χρωματομετρικούς χώρους. Ο CIE-XYZ μπορεί να αναπαρασταθεί μέσω ενός μετασχηματισμού γραμμικών συντεταγμένων του RGB χρωματικού χώρου. Η Y συνιστώσα αντιστοιχεί στη φωτεινότητα του χρώματος. Οι τιμές χρωματικότητας μπορούν να μετασχηματιστούν από το τρισδιάστατο επίπεδο και στη συνέχεια να προβληθούν στο XY. Το αποτέλεσμα είναι το πολύ γνωστό CIE-xy χρωματικό διάγραμμα, το οποίο καθορίζει τον κορεσμό και την απόχρωση του κάθε χρώματος. Ένα από τα μειονεκτήματα των XYZ και xy χρωματικών χώρων, είναι ότι οι διαφορές χρώματος δεν γίνονται εύκολα αντιληπτές σε όλες τις περιοχές του χρωματικού χώρου. Σε αντίθεση ο CIE-lab διαχωρίζει την μεταβλητή της φωτεινότητας L, από τις δυο ενιαίες χρωματικές μεταβλητές a,b. Στην εικόνα 11-h μπορούμε να δούμε την πυκνότητα του χρώματος του δέρματος των Ασιατών στον a,b χρωματικό χώρο. Στην τελευταία σειρά της εικόνας 12 βλέπουμε την πυκνότητα του χρώματος από διαφορετικές φυλές στον a,b χρωματικό χώρο. Παρόλο που οι συγκεκριμένοι χρωματικοί χώροι έχουν πολλά πλεονεκτήματα, χρησιμοποιούνται σπανίως στην ανίχνευση δέρματος, κυρίως επειδή ο μετασχηματισμός από RGB είναι υπολογιστικά "ακριβός".

4.3.2 Ταξινομητές δέρματος

Μια ποικιλία από τεχνικές ταξινόμησης χρησιμοποιούνται στο θεωρητικό κομμάτι για την ταξινόμηση του δέρματος. Ένας ταξινομητής δέρματος είναι ένας πρώτης τάξης ταξινομητής, που καθορίζει ένα όριο απόφασης για την τάξη του χρώματος του δέρματος σε ένα χρωματικό χώρο. Κάθε pixel, του οποίου το χρώμα συμπίπτει με αυτά που υπάρχουν στον ταξινομητή του δέρματος, επισημαίνεται σαν δέρμα. Γι αυτό το λόγο, η επιλογή του ταξινομητή δέρματος είναι άμεσα συνδεδεμένη με το χρωματικό χώρο που θα χρησιμοποιήσουμε. Ο πιο απλός τρόπος για να αποφασίζουμε αν ένα pixel είναι στο χρώμα του δέρματος ή όχι είναι να ορίσουμε ένα όριο. Οι Brand και Mason δημιούργησαν έναν απλό μονοδιάστατο ταξινομητή, σύμφωνα με τον οποίο το pixel ονομαζόταν δέρμα αν η αναλογία μεταξύ των καναλιών R και G ήταν ανάμεσα σε ένα ανώτατο και έναν κατώτατο όριο. Πειραματίστηκαν επίσης με ένα μονοδιάστατο όριο στο IQ επίπεδο του YIQ χώρου όπου το I ήταν η τιμή που χρησιμοποιήθηκε για κατώφλι.

4.3.2.1 Η προσέγγιση του Bayse για ανίχνευση δέρματος

Η ταξινόμηση του δέρματος μπορεί να οριστεί με τη χρήση πιθανοτήτων με τον ακόλουθο τρόπο. Έστω ένα pixel με χρώμα c. Η πιθανότητα να είναι pixel δέρματος είναι $P(\text{skin}|c)$. Όταν υπολογιστεί αυτή η πιθανότητα, το pixel παίρνει την ετικέτα "δέρμα" αν η πιθανότητα αυτή είναι πάνω από το ορισμένο όριο, αλλιώς αν είναι κάτω από το όριο επισημαίνεται σαν μη-δέρμα. Προφανώς όμως, δεν μπορούμε να υπολογίζουμε την πιθανότητα για κάθε πιθανό χρώμα (πχ στο 24 bit RGB υπάρχουν περίπου 1.7 εκατομμύρια χρώματα). Χρησιμοποιώντας τον κανόνα του Bayse ο υπολογισμός της πιθανότητας μπορεί να ξαναγραφεί ως:

$$P(\text{skin}|c) = \frac{P(c|\text{skin})P(\text{skin})}{P(c|\text{skin})P(\text{skin}) + P(c|\text{notskin})P(\text{notskin})}$$

Σύμφωνα με τον κανόνα του Bayse αυτό έχει ως αποτέλεσμα τον υπολογισμό της πιθανής αναλογίας ενός δεδομένου χρώματος σε μια δεδομένη τάξη δέρματος σε σχέση με μια τάξη μη-δέρματος ($P(c|\text{skin})/P(c|\text{non skin})$). Αυτή η αναλογία μπορεί στη συνέχεια να οριστεί σαν κατώφλι για να αποφασιστεί αν ένα pixel είναι δέρμα ή όχι.

4.3.2.2 Εφαρμογές και παραδείγματα ανίχνευσης δέρματος

Ο εντοπισμός και η ανίχνευση του προσώπου είναι το πρώτο βήμα στην απόκτηση βιομετρικών στοιχείων του προσώπου. Το χρώμα του δέρματος είναι ένα χαρακτηριστικό στοιχείο στα ανθρώπινα πρόσωπα. Σε ένα ελεγχόμενο φόντο, η ανίχνευση δέρματος μπορεί να είναι ένα βολικό εργαλείο για τον εντοπισμό προσώπων στην εικόνα. Από τη στιγμή που η επεξεργασία του χρώματος είναι αρκετά πιο γρήγορη από την επεξεργασία οποιουδήποτε άλλου χαρακτηριστικού, η ανίχνευση δέρματος μπορεί να χρησιμοποιηθεί σαν ένα προκαταρκτικό στάδιο για άλλες μεθόδους ανίχνευσης προσώπου. Η ανίχνευση δέρματος χρησιμοποιείται επίσης και για τον εντοπισμό άκρων του σώματος.

Οι Forsyth και Fleck απέδειξαν ότι η ανίχνευση δέρματος μπορεί να χρησιμοποιηθεί ως μέρος της διαδικασίας ανίχνευσης εικόνων με γυμνούς ή προκλητικά ντυμένους ανθρώπους. Η τεχνική τους είχε τρία στάδια. Αρχικά, ένα φίλτρο δέρματος, βασισμένο στο χρώμα και την υφή, χρησιμοποιούνταν για να επιλέξει τις εικόνες με μεγάλες περιοχές στο χρώμα του δέρματος. Έπειτα τα αποτελέσματα, περνούσαν από ένα γεωμετρικό φίλτρο, το

οποίο προσδιόριζε τις περιοχές που ήταν στο χρώμα του δέρματος με κυλινδρικά σχήματα. Αυτά τα κυλινδρικά σχήματα ομαδοποιούνταν σε πιθανά ανθρώπινα άκρα. Οι εικόνες που είχαν μεγάλες ομάδες από άκρα στο χρώμα του δέρματος θεωρούνταν ότι περιείχαν γυμνές φωτογραφίες.

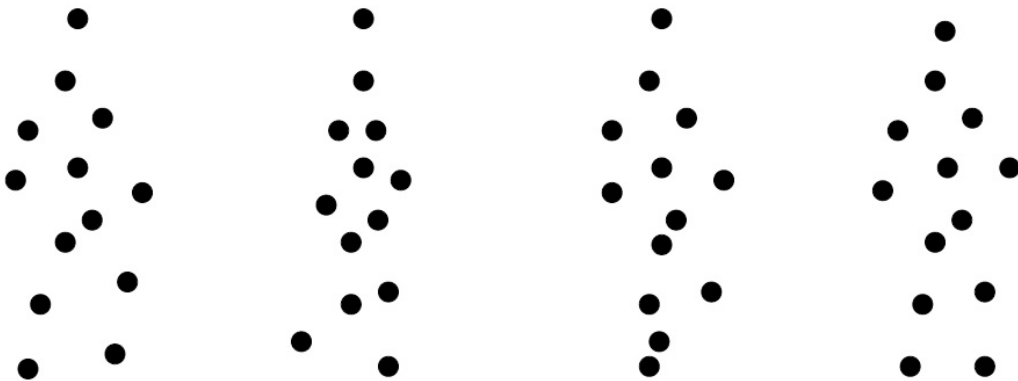
4.3.2.3 Σύνοψη

Ανεξαρτήτως από την επιλογή που θα κάνουμε για το χρωματικό χώρο και τη μέθοδο ταξινόμησης, οι περισσότερες έρευνες πάνω στην ανίχνευση δέρματος έχουν ποσοστό περίπου 60% σωστών ανιχνεύσεων, ενώ το ποσοστό λανθασμένων ανιχνεύσεων κυμαίνεται από 15-30%. Η ανίχνευση δέρματος σε έγχρωμες εικόνες και βίντεο, είναι ένας πολύ εύκολος τρόπος για να ανιχνεύσουμε την παρουσία ανθρώπου. Παρόλα αυτά, αρκετά αντικείμενα στον πραγματικό χώρο έχουν χρώμα στον τόνο του ανθρώπινου δέρματος, τα οποία μπορούν πολύ εύκολα να ανιχνευθούν λανθασμένα ως δέρμα. Γι αυτό το λόγο η ανίχνευση δέρματος έχει εγγυημένα αποτελέσματα μόνο σε ελεγχόμενα περιβάλλοντα στα οποία το φόντο είναι ελεγχόμενο. Επίσης μιας και η ανίχνευση γίνεται με βάση το χρώμα του κάθε pixel η χρήση του περιορίζεται μόνο σε έγχρωμες εικόνες, οπότε δεν έχει καμία χρησιμότητα σε εικόνες gray-scale, υπέρυθρες και γενικότερα όσες δεν έχουν χρωματική πληροφορία.

4.4 Εντοπισμός χαρακτηριστικών σημείων

Η ανίχνευση ανθρώπου βασισμένη μόνο στην κίνηση, αρχικά, φαίνεται εξεζητημένη. Παρέχει η κίνηση των άκρων αρκετή πληροφορία για να εντοπίσει ανθρώπινη παρουσία; Πειράματα που διεξήχθησαν από τον Johansson το 1970 απέδειξαν ότι αρκεί η κίνηση των άκρων προκειμένου να αναγνωριστεί ανθρώπινη παρουσία. Ο Johansson κινηματογράφησε ανθρώπους που κινούντουσαν σε ένα κατάμαυρο δωμάτιο, με μόνη φωτεινή ένδειξη ένα λευκό σημείο φωτός προσκολλημένο στα άκρα τους. Απέδειξε ότι κάποιος που παρακολουθούσε το φιλμ μπορεί εύκολα να αναγνωρίσει την ανθρώπινη κίνηση, παρόλη την απουσία οπτικών ενδείξεων όπως σχήμα, υφή φωτεινότητα και χρώμα. Ένα παράδειγμα των σημείων του Johansson φαίνεται στην εικόνα 17.

Δεδομένου ότι ο άνθρωπος μπορεί χωρίς προσπάθεια να αναγνωρίσει την κίνηση, είναι πιθανό και ένας αλγόριθμος να κάνει το ίδιο. Επιπρόσθετα, τα μεμονωμένα σημεία κίνησης όπως χρησιμοποιούνται στο πείραμα του Johansson, μπορούν να αναπαρασταθούν αποτελεσματικά σε υπολογιστή. Αυτό μας δίνει την ελπίδα ότι ένας αποδοτικός αλγόριθμος είναι πιθανό να επιτευχθεί .



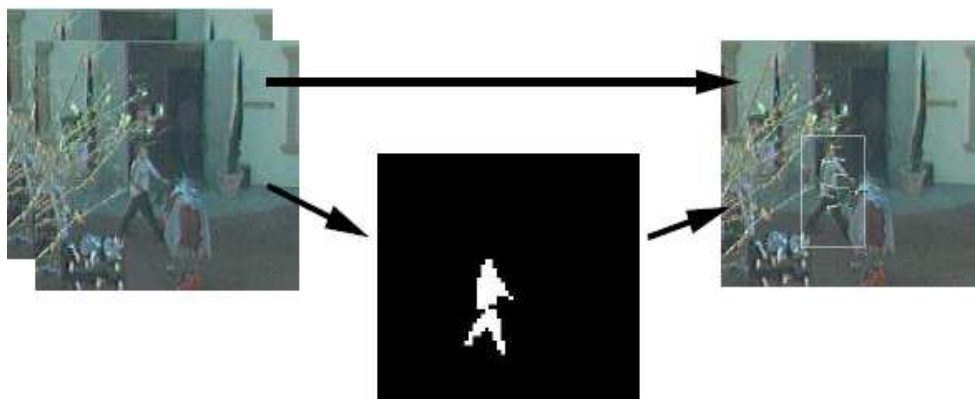
Εικόνα 18: Ακολουθία από σημεία του Johansson που δείχνει την πλαϊνή λήψη ενός ανθρώπου που περπατάει.

Το πείραμα του Johansson μας δείχνει ότι η ανίχνευση ανθρώπου από την κίνηση των χαρακτηριστικών αυτών σημείων είναι ένας ρεαλιστικός στόχος. Με αυτό το σκεπτικό στο μυαλό μας, μπορούμε να χωρίσουμε ένα τέτοιο σύστημα σε 3 διαφορετικές ρουτίνες.

- I. Εξαγωγή των χαρακτηριστικών σημείων από βίντεο.
- II. Εύρεση ενός μοντέλου που αντιπροσωπεύει με ακρίβεια την ανθρώπινη κίνηση
- III. Εφαρμογή αυτού του μοντέλου σε έναν ανιχνευτή, ο οποίος θα αποφασίζει αν μια ομάδα από κινούμενα σημεία είναι αντιπροσωπευτική ανθρώπινης κίνησης.

4.4.1 Κατανόηση του προβλήματος

Η προσέγγιση της Song στην ανίχνευση ανθρώπινης κίνησης είναι ένα πιθανολογικό μοντέλο, στο οποίο τα κοινά χαρακτηριστικά και μεταβλητές της ανθρώπινης κίνησης εισάγονται σε ένα μοναδικό στατιστικό μοντέλο. Αυτό το μοντέλο αποθηκεύεται σαν ένα γράφημα το οποίο περιέχει ακμές και γωνίες. Κάθε κορυφή αναπαριστά ένα χαρακτηριστικό κίνησης σε κάποιο σημείο από ανθρώπινη σιλουέτα. Η κορυφή αναπαριστάται από μια ευκλείδεια θέση (x,y) η οποία προσδιορίζει το μέσο όρο της θέσης των χαρακτηριστικών αυτών, στον ανθρώπινο σώμα αλλά και το μέσο όρο της ταχύτητας με την οποία κινούνται (U_x, U_y) . Επιπρόσθετα, υπάρχει άλλος ένας πίνακας ο οποίος συσχετίζει αυτές τις τέσσερις παραμέτρους με αυτές σε κάθε άλλη κορυφή. Κάθε 2 κορυφές οι οποίες συνδέονται με μια ακμή, στατιστικώς θεωρούνται εξαρτημένες η μια από την άλλη, και αυτές που δεν ενώνονται θεωρούνται ανεξάρτητες. Σε αυτό το σημείο πρέπει να αναφέρουμε ότι η αποδοτικότητα ενός τέτοιου συστήματος είναι $O(MN^3)$, όπου M είναι ο αριθμός των τριγώνων σε ένα γράφημα και N ο αριθμός των σημείων που αξιολογούμε. Γι αυτό το λόγο, πρέπει να φροντίσουμε να κρατήσουμε αυτά τα σημεία στο ελάχιστο. Η μείωση των σημείων αυτών απαιτεί πιο έξυπνη επιλογή τους. Θα ήταν μη πρακτικό να βρούμε κάθε σημείο κίνησης σε όλη την εικόνα και να τα εισάγουμε στο σύστημα για αξιολόγηση. Αρχικά τα σημεία αυτά θα ήταν παρά πολλά με αποτέλεσμα να έχουμε πολύ αργή απόδοση. Δεύτερον, θα δίναμε άχρηστα σημεία όπως είναι αυτά που ανήκουν στο background. Για να αποφύγουμε τη σπατάλη χρόνου εξετάζοντας το background, χρησιμοποιούμε image segmentation για να διαχωρίσουμε το πίσω από το μπροστά μέρος μιας σκηνής. Η κατάτμηση εικόνας είναι η ενέργεια που κάνει οποιοσδήποτε αλγόριθμος ο οποίος διαχωρίζει τις περιοχές μιας εικόνας με τέτοιο τρόπο, όπως θα τις διαχώριζε και ένας άνθρωπος. Μιας και μας ενδιαφέρει η κίνηση, η φυσική προσέγγιση είναι να απομονώσουμε τις περιοχές της εικόνας που κινούνται, από το background. Αφού διαχωρίσουμε την εικόνα με βάση την κίνηση, τα σημεία που μας ενδιαφέρουν μπορούν να εντοπιστούν ξεχωριστά σε κάθε περιοχή. Αυτό μας επιτρέπει να τρέξουμε την αξιολόγηση ξεχωριστά για κάθε περιοχή.



Εικόνα 19: Κατάτμηση Εικόνας

Αφού η αξιολόγηση του μοντέλου έχει τύπο $O(MN^3)$ τρέχει πιο γρήγορα σε 2 περιοχές με 10 σημεία ενδιαφέροντος παρά σε μια περιοχή 40 σημείων. Αυτά τα δυο βασικά στάδια, η κατάτμηση εικόνας και ο εντοπισμός σημείων συνθέτουν το πρώτο μισό από ένα σύστημα ανίχνευσης κίνησης

4.4.2 Κατάτμηση εικόνας

Ο στόχος μας στην κατάτμηση εικόνας, είναι να διαχωρίσουμε την περιοχή του background με αυτή του foreground που παρουσιάζει ενδιαφέρον, ως προς την ανίχνευση κίνησης. Σε ιδανικές συνθήκες, δηλαδή αν τα στοιχεία του background δεν μετακινούνται και δεν υπάρχει απότομη αλλαγή στην φωτεινότητα του, ο αλγόριθμος, ο οποίος κάνει την αφαίρεση του background, δημιουργεί ένα μοντέλο της εικόνας του φόντου με την πάροδο του χρόνου. Για κάθε frame του video αφαιρούμε το background του frame από αυτό του μοντέλου. Τα pixel με αποτέλεσμα κοντά στο 0 αντιμετωπίζονται σαν φόντο και αυτά με μεγαλύτερη τιμή σαν προσκήνιο. Έτσι λοιπόν, όταν έχουμε το μοντέλο του background ο αλγόριθμος είναι απλός και εύκολος στην εισαγωγή του. Το να αποκτήσουμε όμως το μοντέλο του φόντου είναι αρκετά περίπλοκο. Η πιο εύκολη προσέγγιση είναι να στήσουμε την κάμερα, να αδειάσουμε τη σκηνή από κινούμενα αντικείμενα και να πάρουμε έτσι ένα στιγμιότυπο. Κάτι τέτοιο όμως δεν είναι εφαρμόσιμο, σε εξωτερικές και πραγματικές σκηνές καθώς το φόντο μπορεί να αλλάξει με την πάροδο του χρόνου ή να είναι δύσκολο να αδειάσουμε τη σκηνή. Ο φωτισμός μπορεί να αλλάξει απότομα και η κάμερα να μετακινηθεί. Μια πιο πρακτική προσέγγιση είναι μια που μπορεί να προσαρμοστεί σε ένα ήπια μεταβαλλόμενο φόντο, το οποίο θα περιγράψουμε παρακάτω. Ας υποθέσουμε ένα χρονικά μεταβαλλόμενο pixel στη θέση (x,y) σε μια grayscale ακολουθία video. Θα αναφερόμαστε σε αυτή την τιμή ως $V_{x,y}(t)$. Μπορούμε να συμπεριφερθούμε στην τιμή σαν μια τυχαία συνάρτηση του X_t

$$X_t = V_{x,y}(t)$$

Τώρα υποθέτουμε ότι μπορούμε να μοντελοποιήσουμε την πιθανότητα παρατήρησης της τρέχουσας τιμής του Pixel ως μια μίξη από Gaussian κατανομές. Αυτή η πιθανότητα είναι,

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} \eta(X_t, \mu_{i,t}, \Sigma_{i,t})$$

Εφόσον το φόντο υποθέτουμε ότι είναι στατικό, η αξία των pixel τα οποία αποτελούν το φόντο μπορούν να αναπαρασταθούν με ένα ή περισσότερα Gaussians με μικρή μεταβολή εξαιτίας του θορύβου της εικόνας. Περισσότερα από ένα Gaussian είναι πιθανά σε σκηνές όπως δέντρων που κουνιούνται από τον αέρα ή φωτός που αναβοσβήνει. Επιπλέον, στις περισσότερες σκηνές, το φόντο μπορεί να είναι ορατό, συχνότερα απ' ό,τι το προσκήνιο σε συγκεκριμένα pixel, οπότε το Gaussian με το μεγαλύτερο βάρος ω είναι πιθανά το background. Αυτές οι ιδέες μας επιτρέπουν να προσεγγίσουμε το πρόβλημα της αφαίρεσης background με την ακόλουθη μεθοδολογία

- Για κάθε pixel σε ένα καρέ βίντεο:
 - Λαμβάνουμε υπόψη μας τις τελευταίες N τιμές που έχει πάρει το συγκεκριμένο pixel.
 - Βρίσκουμε τα K Gaussians και τα βάρη που ταιριάζουν καλύτερα σε αυτό το δείγμα από N τιμές χρησιμοποιώντας συγκεκριμένους αλγόριθμους όπως οι K -Means ή Expectation Maximization (EM)

- Επιλέγουμε το Gaussian με το μεγαλύτερο βάρος ω και αποθηκεύουμε το μέσο όρο, σαν την τιμή της εικόνας του φόντου για το συγκεκριμένο pixel.
- Αφαιρούμε το φόντο από το καρέ.
- Στην προκύπτουσα εικόνα, οποιαδήποτε τιμή μεγαλύτερη από τρεις τυπικές αποκλίσεις από το μέσο όρο, θεωρείται foreground, και οποιαδήποτε άλλη τιμή θεωρείται φόντο.

Ο αλγόριθμος που προηγήθηκε είναι υπερβολικά απαιτητικός σε υπολογιστική ισχύ και αυτό καθιστά τη χρήση του σε πραγματικό χρόνο αρκετά δύσκολη, ιδίως το βήμα της τοποθέτησης των K Gaussians στον πίνακα δεδομένων του κάθε pixel για κάθε καρέ. Για την απλούστευση του συστήματος, η εικόνα του φόντου πρέπει να επαναυπολογίζεται μόνο κάθε N frame. Έτσι, για τα περισσότερα χρονικά βήματα, οι τιμές του κάθε pixel απλά συλλέγονται και αποθηκεύονται για μετέπειτα επεξεργασία κάτι που γίνεται μία φορά ανά N frame. Το μειονέκτημα αυτής της προσέγγισης είναι ότι υπάρχει μια χρονική καθυστέρηση, για να μπορέσει ο αλγόριθμος να προσαρμόσει στο φόντο καινούργια ακίνητα αντικείμενα.

4.4.3 Εντοπισμός χαρακτηριστικών σημείων

Το βήμα της κατάτμησης εικόνας μας επιτρέπει να διαχωρίζουμε τα αντικείμενα στο προσκήνιο από το φόντο. Παρόλα αυτά, συνεχίζουμε να δουλεύουμε πάνω σε ολόκληρη την εικόνα και όχι στα επιμέρους σημεία κίνησης τα οποία είναι επιθυμητά για την ανίχνευση ανθρώπινης κίνησης. Το πρόβλημα του υπολογισμού της κίνησης σε μια εικόνα είναι γνωστό ως εντοπισμός της οπτικής ροής της εικόνας. Υπάρχει μια μεγάλη ποικιλία από τεχνικές για τον υπολογισμό της οπτικής ροής, αλλά η μέθοδος Kanade-Lucas-Tomasi ξεχωρίζει για την απλότητά της, αλλά και την απουσία παραδοχών σχετικά με την υποκείμενη εικόνα. Ένα απλό παράδειγμα του αλγόριθμου αυτού φαίνεται στο σχήμα X. Ο πιο απλοϊκός αλγόριθμος για να επιτύχουμε point feature tracking σε 2 καρέ βίντεο παρουσιάζεται στον παρακάτω πίνακα(πίνακας X)

- Διαλέγουμε ένα μικρό παράθυρο, για παράδειγμα 7 pixel σε κάθε κατεύθυνση, γύρω από ένα pixel που παρουσιάζει ενδιαφέρον στο πρώτο καρέ. Αυτό το pixel το ονομάζουμε pixel A.
- Κάθε pixel κοντά στο A, στο δεύτερο καρέ, το ονομάζουμε pixel B
- Αφαιρούμε την τιμή του κάθε pixel στην 7x7 περιοχή γύρω από το pixel A από κάθε pixel της 7x7 περιοχής γύρω από το pixel B. Υψώνουμε στο τετράγωνο το αποτέλεσμα της διαφοράς, και αθροίζουμε αυτές τις 49 τιμές για να παράγουμε την 'ανομοιότητα' για την επιλογή αυτού του pixel B.
- Το pixel B στο δεύτερο καρέ με τη μικρότερη τιμή ανομοιότητας θεωρείται ότι είναι η νέα θέση του pixel A.

Παρά το γεγονός ότι ο αλγόριθμος αυτός θα μας δώσει μια νέα θέση και ταχύτητα για το pixel A, έχει κάποιες σημαντικές αδυναμίες. Αρχικά, θα είναι αρκετά αργός, αφού θα απαιτεί εκατοντάδες υπολογισμούς σε κάθε επανάληψη και εκατοντάδες επαναλήψεις. Δεύτερον, ο αλγόριθμος θα μας δίνει την θέση και την ταχύτητα του σημείου μόνο στο κοντινότερο ακέραιο pixel. Αν το σημείο έχει κινηθεί ενάμιση pixel θα πρέπει να το υπολογίσουμε, είτε στη θέση 1, είτε στη θέση 2. Ο αλγόριθμος Kanade-Lucas-Tomasi μειώνει αυτά τα προβλήματα με την επαναλαμβανόμενη χρήση των βαθμίδων χρώματος της

εικόνας, για να προβλέψει τη νέα θέση του σημείου, μέχρι αυτή να επιτευχθεί. Επειδή αυτή η προσέγγιση είναι βασισμένη στην σειρά επέκτασης Taylor (Taylor series expansion) δεν κάνει υποθέσεις σχετικά με την εικόνα που βρίσκεται από κάτω. Ένα τελευταίο χαρακτηριστικό του αλγόριθμου Kanade-Lucas-Tomasi, είναι η επιλογή των αρχικών σημείων. Είναι χάσιμο χρόνου να προσπαθήσουμε να ανιχνεύσουμε τα ίχνη όλων των pixel από την αρχική εικόνα μέχρι την εικόνα προορισμού. Μια πιο χρήσιμη προσέγγιση είναι να ανιχνεύσουμε μόνο τα pixel τα οποία αναπαριστούν έντονα, σαφώς διακρινόμενα σημεία. Γι αυτό το λόγο, όταν διαλέγουμε τα σημεία που θα ανιχνεύσουμε, ταξινομούμε τα pixel έτσι ώστε οι ιδιοτιμές τους να είναι σε φθίνουσα σειρά και διαλέγουμε τις N πρώτες, όπου N ο αριθμός των pixel που θέλουμε να ανιχνεύσουμε.

4.4.4 Ανίχνευση Μοντέλου Κίνησης

Έστω ένα σετ από κινούμενα σημεία, κάθε ένα με μια θέση και ταχύτητα, ο σκοπός μας στη ανίχνευση μοντέλου κίνησης είναι να αποφασίσουμε αν μια ομάδα από σημεία είναι αντιπροσωπευτική ενός προκαθορισμένου μοντέλου. Η προσέγγιση της Y. Song για την αποθήκευση του μοντέλου σαν μια συνάρτηση πιθανοτήτων, και η εύρεση μιας επισήμανσης για τα δεδομένα που μεγιστοποιούν την πιθανότητα, είναι μια από της πιο γνωστές και θα την περιγράψουμε στη συνέχεια του κεφαλαίου. Αρχικά ορίζουμε μια ομάδα από ανθρώπινα μέλη, το καθένα από τα οποία αντιστοιχεί σε ένα χαρακτηριστικό σημείο, το οποίο θα μπορούσαμε να παρακολουθήσουμε. Έστω λοιπόν $S_{body} = \{AK, AA, A\Omega, K, \dots, \Delta\Pi\}$ το σετ των μελών του σώματος, όπου AK ο αριστερός Καρπός Αγκώνας Ωμος, αντίστοιχα, K το κεφάλι, Π το πόδι κ.ο.κ. Θα αναφερόμαστε σε αυτά σαν το σετ των πιθανών ετικετών. Από τη στιγμή που το μοντέλο για τον άνθρωπο μαθαίνεται χωρίς επίβλεψη, η πραγματική χαρτογράφηση των μελών του σώματος σε χαρακτηριστικά σημεία είναι απροσδιόριστη και δεν αντιστοιχεί επακριβώς σε μεμονωμένα άκρα. Κάθε μέλος του σώματος έχει ένα διάνυσμα από μετρήσεις, το οποίο αποτελείται από τη θέση και την ταχύτητα, και στο οποίο θα αναφερόμαστε σαν X_{AK}, X_{AA} κτλ. Το μοντέλο κίνησης αποθηκεύεται σαν μια συνάρτηση πιθανοτήτων P , που μπορεί να αξιολογηθεί για ένα δεδομένο σύνολο δεδομένων.

$$P_{S_{body}}(X_{AK}, X_{AA}, X_{A\Omega}, X_K, \dots, X_{\Delta\Pi}) \quad (2.3)$$

Τώρα, θεωρούμε ένα διάνυσμα από N παρατηρούμενα σημεία $\overline{X} = [X_1, X_2, \dots, X_N]$ στο οποίο η αντιστοιχία μεταξύ σημείων και ετικετών δεν είναι γνωστή εκ των προτέρων. Θέλουμε να βρούμε την μετάθεση των σημείων που μεγιστοποιούν τη συνάρτηση της πιθανότητας, (εξίσωση 2.3). Με άλλα λόγια, θέλουμε να βρούμε μια επισήμανση $\overline{L} = [L_1, L_2, \dots, L_N]$ όπου $L_i \in S_{body}$ είναι η ετικέτα του σημείου X_i , που μεγιστοποιεί τη συνάρτηση της πιθανότητας. Θα καθορίσουμε αυτή την βέλτιστη επισήμανση σαν:

$$\overline{L} = \operatorname{argmax} P(\overline{L} | \overline{X})$$

Όπου $P(\overline{L} | \overline{X})$ η πιθανότητα παρατήρησης του \overline{X} δεδομένης της επισήμανσης \overline{L} . Ας υποθέσουμε ότι N ο αριθμός των σημείων που ισούνται με τα μέλη του σώματος, η προσπάθεια για τη λύση της εξίσωσης αυτής με brute force θα είχε ως αποτέλεσμα χρόνο εκτέλεσης $O(e^N)$. Μία τέτοια κακή απόδοση όμως είναι απαγορευτική για ένα πρακτικό αλγόριθμο. Το βασικό επίτευγμα, της προσέγγισης της Song είναι η υπόθεση ότι ορισμένες παράμετροι της συνάρτησης πιθανότητας είναι ανεξάρτητες μεταξύ τους. Με τον καθορισμό μιας τέτοιας δομής κατά την εκμάθηση του μοντέλου, η αποτελεσματικότητα του ανιχνευτή μπορεί να βελτιωθεί πάρα πολύ χρησιμοποιώντας ένα δυναμικό αλγόριθμο προγραμματισμού. Πριν κατανοήσουμε γιατί μια τέτοια βελτίωση είναι εφικτή, πρέπει πρώτα να καταλάβουμε πως αναπαριστάται το μοντέλο. Όπως περιγράψαμε, το μοντέλο κίνησης αποθηκεύεται σαν ένα γράφημα στο οποίο κάθε κορυφή αναπαριστάται από ένα χαρακτηριστικό σημείο, και κάθε ακμή αναπαριστά μια στατιστική εξάρτηση ενός

χαρακτηριστικού από ένα άλλο. Μια βοηθητική υπόθεση είναι να αναγκάσουμε το γράφημα να έχει μια ειδική μορφή ενός αποσυντιθέμενου τριγωνικού γραφήματος. Εξαιτίας αυτής της δομής του, το γράφημα θα έχει μια κορυφή η οποία εξαρτάται από μόνο 2 κορυφές. Όταν εξαιρεθεί αυτή η κορυφή, θα υπάρξει κάποια άλλη κορυφή η οποία εξαρτάται από δυο άλλες και ούτω καθεξής. Η αλληλουχία της αφαίρεσης κορυφών είναι γνωστή σαν σειρά εξάλειψης του γραφήματος. Στο σχήμα X βλέπουμε ένα απλό παράδειγμα ενός τέτοιου γραφήματος και της σειράς εξάλειψης κορυφών. Επομένως, μια συνάρτηση πιθανοτήτων η οποία είναι αποθηκευμένη σε έναν αποσυντιθέμενο τριγωνικό γράφημα, μπορεί να προσεγγιστεί σαν ένα προϊόν ανεξάρτητων συναρτήσεων. Για παράδειγμα, αν η σειρά εξάλειψης των κορυφών του γραφήματος είναι A,B,C,D,E,F, όπως είναι στο σχήμα μας η συνάρτηση πιθανότητας μπορεί να μετασχηματιστεί σε:

$$P(A,B,C,D,E,F)=P(A|B,E)P(B|E,F)P(C|E,F)P(D,E,F)$$

Επειδή, η συνάρτηση της πιθανότητας με πολλές παραμέτρους, μπορεί πλέον να αντιπροσωπευτεί από ένα προϊόν μικρότερων εξαρτημένων συναρτήσεων, ένας αλγόριθμος δυναμικού προγραμματισμού είναι εφικτός, και μπορεί να ψάξει για όλες τις πιθανές επισημάνσεις πολύ πιο βολικά. Κάθε εξαρτημένη συνάρτηση, μπορεί να ανιχνευθεί ατομικά στο $O(N^3)$ προκειμένου να εντοπίσει την ετικέτα που την μεγιστοποιεί για κάθε δυνατό ζεύγος ετικετών. Επομένως, για ένα μοντέλο με M κορυφές, ο συνολικός χρόνος εκτέλεσης προκειμένου να μεγιστοποιηθεί η συνάρτηση της πιθανότητας είναι $O(M N^3)$, το οποίο αποτελεί μια πολύ σημαντική βελτίωση από το $O(e^N)$ που προκύπτει από την τεχνική brute force. Με αυτό τον τρόπο, μπορούμε να βρούμε την ετικέτα \overline{L} που μεγιστοποιεί την συνάρτηση πιθανότητας P . Για ένα δοσμένο \overline{X} , αυτή η διαδικασία μας δίνει δυο χρήσιμες πληροφορίες. Αρχικά την αριθμητική τιμή της συνάρτησης πιθανοτήτων και δεύτερον την βέλτιστη επισήμανση των χαρακτηριστικών σημείων. Προκειμένου να ολοκληρώσουμε την διαδικασία της ανίχνευσης, πρέπει να διαλέξουμε ένα κατώφλι για την τιμή της συνάρτησης πιθανοτήτων. Αν η τιμή είναι μεγαλύτερη από το κατώφλι, θεωρούμε ότι τα σημεία που δώσαμε σαν είσοδο ταιριάζουν στο μοντέλο, ενώ αν είναι μικρότερη λέμε ότι δεν υπήρξε ταύτιση. Η τιμή του κατωφλιού διαλέγεται εμπειρικά σαν συνάρτηση των λάθος ανιχνεύσεων και των ανιχνεύσεων που δεν βρέθηκαν.

5 Υλοποίηση

Σε αυτό το σημείο θα περάσουμε στο πρακτικό μέρος της πτυχιακής παρουσιάζοντας το εργαλείο με το οποίο θα υλοποιήσουμε την πτυχιακή μας. Στη συνέχεια θα δούμε ένα διάγραμμα, σύμφωνα με το οποίο θα πραγματοποιήσουμε το πρόγραμμα, το οποίο θα ανιχνεύει ανθρώπινη κίνηση στο χώρο.

5.1 Εισαγωγή στο Matlab

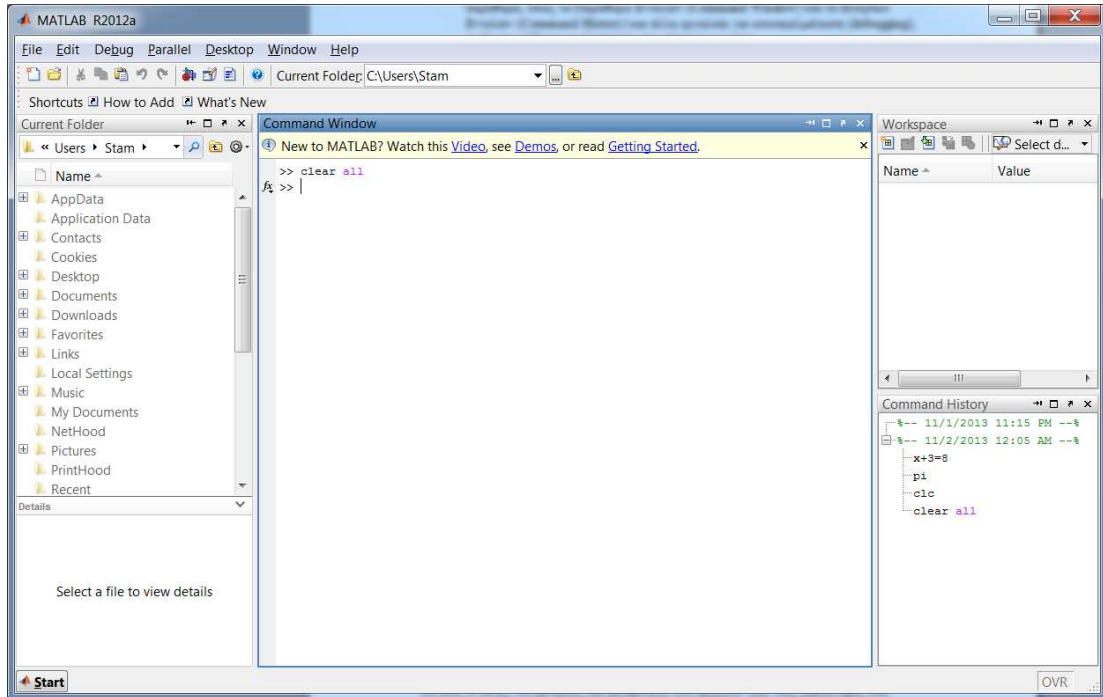
Το Matlab είναι ένα εργαλείο το οποίο προσφέρει ένα διαδραστικό προγραμματιστικό περιβάλλον στον χρήστη και χρησιμοποιείται σε ένα μεγάλο αριθμό εφαρμογών. Στους χρήστες του Matlab προσφέρεται μια υψηλού επιπέδου γλώσσα προγραμματισμού, κατάλληλη για τη μοντελοποίηση και επίλυση σύνθετων μαθηματικών, και όχι μόνο, προβλημάτων.

Το όνομά του προέρχεται από τις λέξεις Matrix Laboratory (εργαστήριο πινάκων) επειδή για τη λειτουργία του χρησιμοποιεί εξολοκλήρου πίνακες.

Το Matlab εκτός των άλλων, επιτρέπει τον εύκολο χειρισμό πινάκων, τη γραφική απεικόνιση (plotting) συναρτήσεων και δεδομένων, την υλοποίηση αλγορίθμων και τη συνεργασία και διαλειτουργικότητα με προγράμματα γραμμένα σε άλλες γλώσσες προγραμματισμού. Επιπροσθέτως επειδή το Matlab μπορεί να χρησιμοποιηθεί σε πολλά επιστημονικά πεδία(επεξεργασία σήματος, νευρωνικά δίκτυα, συστήματα ελέγχου κ.τ.λ.), είναι δυνατή η ενσωμάτωση πρόσθετων πακέτων, που ονομάζονται toolboxes, και δίνουν τη δυνατότητα χρήσης εξειδικευμένων συναρτήσεων. Τα βασικά συστατικά του Matlab είναι:

- το Περιβάλλον Ανάπτυξης και τα αντίστοιχα εργαλεία του: περιλαμβάνει διάφορα παράθυρα, όπως το Παράθυρο Εντολών (Command Window) και το Ιστορικό Εντολών (Command History) και άλλα εργαλεία για αποσφαλμάτωση (debugging), ανάλυση κώδικα και πλοήγηση στο σύστημα αρχείων
- η βιβλιοθήκη μαθηματικών συναρτήσεων: ένα από τα πιο σημαντικά συστατικά του Matlab, το οποίο προσφέρει μεγάλο εύρος αριθμητικών συναρτήσεων, από τις πιο απλές μέχρι τις πιο περίπλοκες.
- η γλώσσα προγραμματισμού: μια υψηλού επιπέδου προγραμματιστική γλώσσα με δομές δεδομένων, συναρτήσεις, εντολές ελέγχου ροής, εντολές εισόδου/εξόδου και στοιχεία από αντικειμενοστραφείς γλώσσες προγραμματισμού
- τα γραφικά συστατικά: το Matlab παρέχει μια πληθώρα δυνατοτήτων απεικόνισης διανυσμάτων, πινάκων και γραφημάτων στις 2 και 3 διαστάσεις.

Στο σημείο αυτό θα περιγράψουμε το περιβάλλον ανάπτυξης (Development Environment) του Matlab.



Εικόνα 20: Matlab Desktop (Επιφάνεια Εργασίας Matlab)

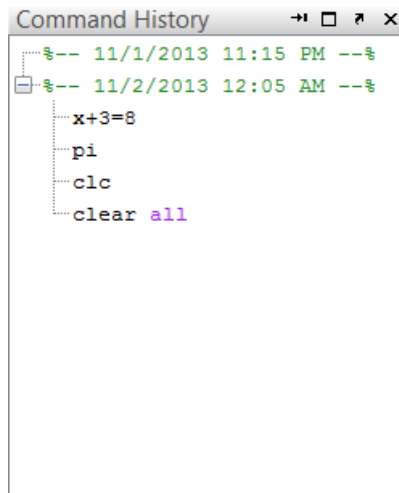
Το MATLAB Desktop (Επιφάνεια Εργασίας MATLAB) αποτελεί τη βάση του Περιβάλλοντος Ανάπτυξης του συγκεκριμένου εργαλείου. Παρέχει στο χρήστη ένα προγραμματιστικό περιβάλλον στο οποίο του δίνεται η δυνατότητα να εκτελέσει εντολές, να αστικοποιήσει δεδομένα και αποτελέσματα, να δημιουργήσει και να τροποποιήσει αρχεία m-files (απλά αρχεία κειμένου με κατάληξη .m που περιέχουν εντολές MATLAB) και να λάβει στοχευόμενη βοήθεια. Αποτελείται από έναν αριθμό άλλων παραθύρων των οποίων η χρησιμότητα θα αναλυθεί στην συνέχεια.



Εικόνα 21: Παράθυρο Εντολών (Command Window)

Πτυχιακή Εργασία Τμήματος Μηχανικών Πληροφορικής

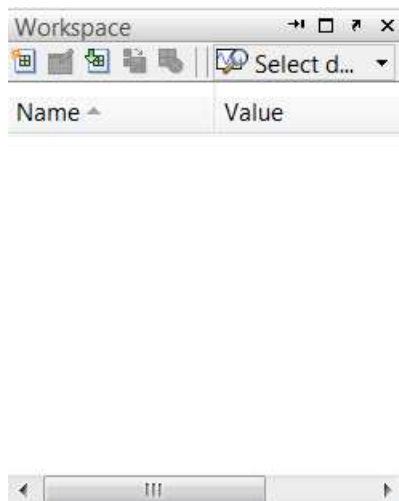
Στο παράθυρο εντολών μπορούμε να εκτελέσουμε διάφορες εντολές σε γλώσσα Matlab, προκειμένου να ολοκληρώσουμε έναν αριθμό εργασιών, όπως για παράδειγμα την δημιουργία και επεξεργασία μεταβλητών (variables), τη δημιουργία, τροποποίηση και κλήση αρχείων m-files, τη διαχείριση δεδομένων και πολλά άλλα. Στο παράθυρο αυτό η εκτέλεση της εντολής πραγματοποιείται με τη χρήση του πλήκτρου «Enter» και τα αποτελέσματα της εντολής γίνονται άμεσα εμφανή κάτω από την εντολή μας. Να επισημάνουμε σε αυτό το σημείο ότι με τη χρήση του πλήκτρου «Tab» εμφανίζεται μια λίστα βοήθειας με όλες τις δυνατές εντολές, συναρτήσεις και μεταβλητές που αρχίζουν από τους χαρακτήρες που έχουμε εισάγει. Επίσης, είναι δυνατή η επανάκληση προηγούμενα εκτελεσμένων εντολών απλά χρησιμοποιώντας τα Up & Down Arrow Keys (πάνω και κάτω βελάκι).



```
Command History
-- 11/1/2013 11:15 PM --
x+3=8
pi
clc
clear all
```

Εικόνα 22: Ιστορικό Εντολών (Command History)

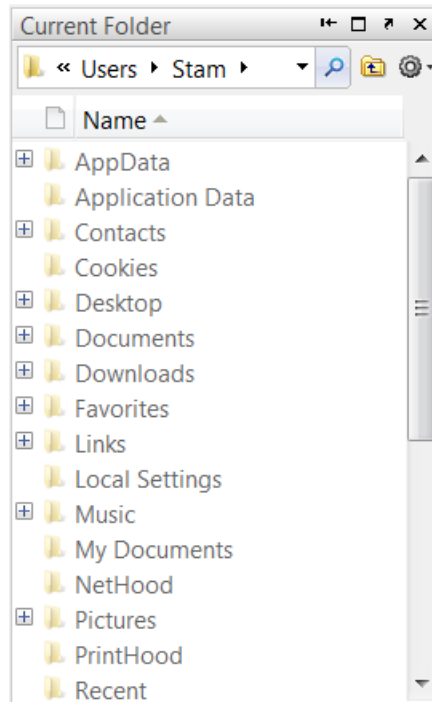
Όλες οι εντολές που έχουμε εκτελέσει στο παράθυρο εντολών (Command Window) αποθηκεύονται στο παράθυρο αυτό. Οι εγγραφές αυτές, ταξινομούνται σύμφωνα με την ώρα εκτέλεσής τους. Μπορούμε με τη βοήθεια αυτού του παραθύρου να εκτελέσουμε ξανά εντολές που έχουμε ήδη εισάγει, ή να τις επεξεργαστούμε περισσότερο σέρνοντας τις στο παράθυρο εντολών. Οι εγγραφές του Ιστορικού Εντολών, τέλος, μπορούν να διαγραφούν, να ομαδοποιηθούν, να τροποποιηθούν ή να σωθούν σε ένα m-file.



Εικόνα 23: Χώρος εργασίας (Workspace)

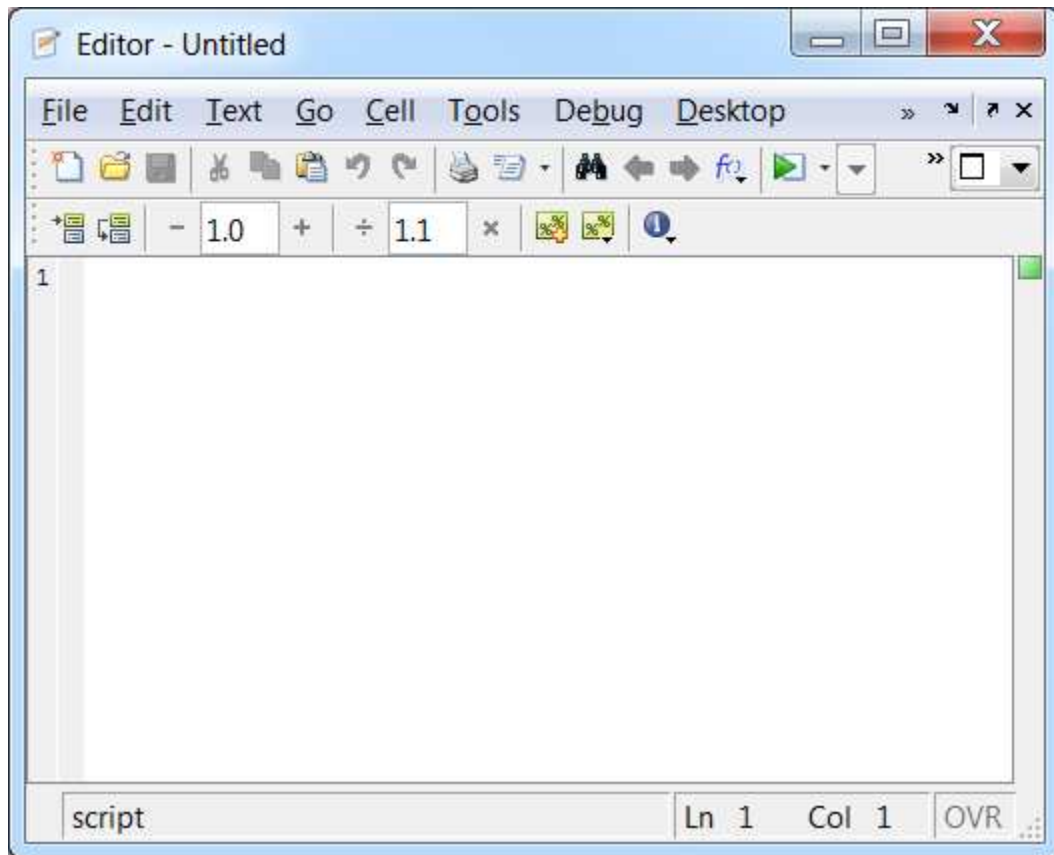
Ο χώρος εργασίας αποτελείται από όλες τις μεταβλητές που έχουν οριστεί στο παράθυρο εντολών και είναι προσβάσιμες από όλα τα προγράμματα. Κάθε μεταβλητή που αποθηκεύεται έχει διάφορες πληροφορίες όπως η τιμή της, το όνομά της, διάφορα στατιστικά

στοιχεία και άλλα. Οι πληροφορίες αυτές γίνονται εμφανείς στον χρήστη μέσω του Workspace Browser. Δυνατότητες όπως η διαδραστική ανάγνωση και τροποποίηση των χαρακτηριστικών των αποθηκευμένων μεταβλητών (με διπλό-κλικ επάνω στην επιθυμητή μεταβλητή και χρήση του Array Editor) και η γραφική απεικόνισή τους (plotting) είναι μερικές από τις παρεχόμενες από αυτόν ευκολίες. Η συμβολή του, δε, κατά την διαδικασία της αποσφαλμάτωσης (debugging) προγραμμάτων είναι καταλυτική.



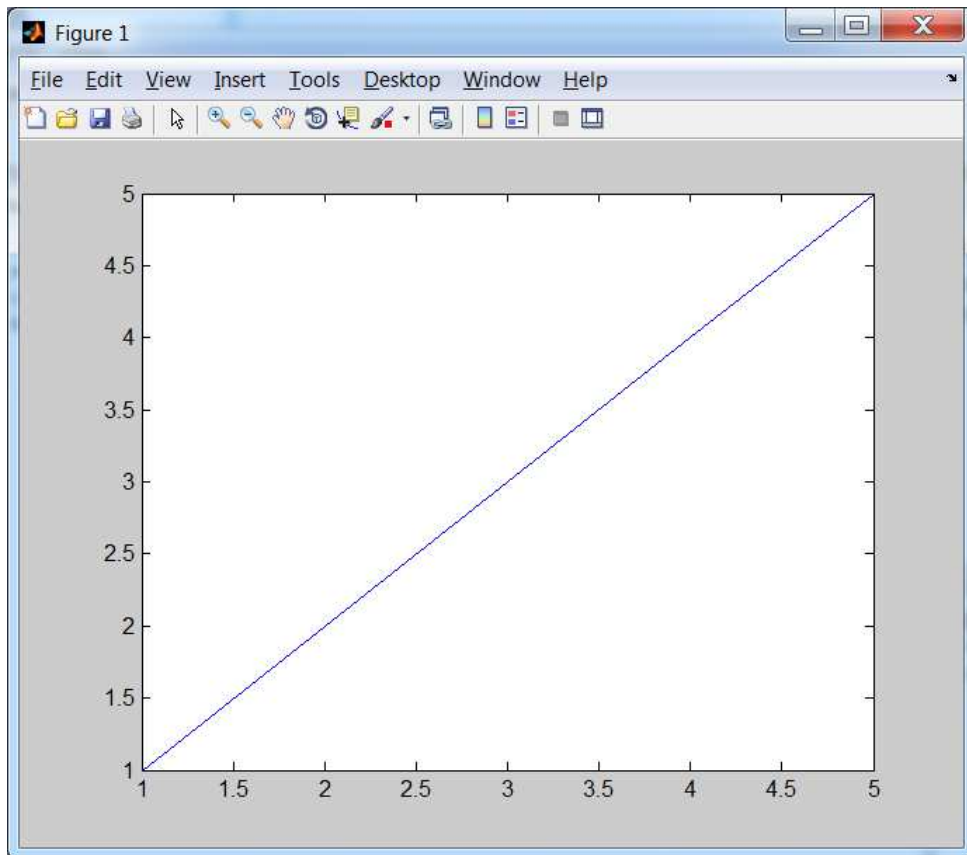
Εικόνα 24: Τρέχων Κατάλογος (Current Directory)

Κάθε λειτουργία η οποία στηρίζεται στην χρήση αρχείων, χρησιμοποιεί τον Τρέχοντα Κατάλογο (Current Directory) ως σημείο αναφοράς. Κάθε αρχείο που επιθυμούμε να εκτελέσουμε ή να χρησιμοποιήσουμε θα πρέπει αναγκαστικά να βρίσκεται μέσα στο συγκεκριμένο φάκελο εκτός και αν εμείς το ορίσουμε διαφορετικά. Στον κατάλογο αυτό παρουσιάζονται τα περιεχόμενα του, επιτρέποντας την περιήγηση του χρήστη στο σύστημα αρχείων του εκάστοτε υπολογιστή, το άνοιγμα (με το κατάλληλο προκαθορισμένο πρόγραμμα), την αναζήτηση και τροποποίηση σχετικών αρχείων ή καταλόγων καθώς και την διαχείριση m-files.



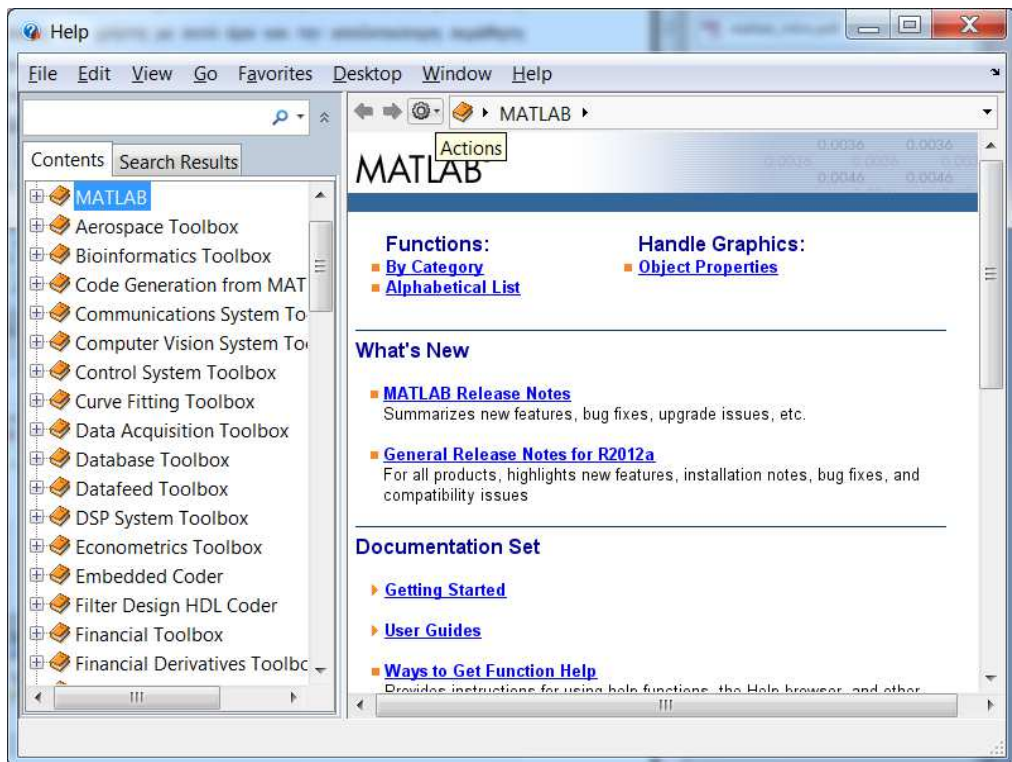
Εικόνα 25: Παράθυρο Εισαγωγής Κειμένου (Editor window)

Στο παράθυρο αυτό μπορούμε να δημιουργήσουμε μια ακολουθία εντολών και συναρτήσεων και να τα αποθηκεύσουμε ως m-file, έτσι ώστε να μπορέσουμε να τα εκτελέσουμε οποιαδήποτε στιγμή. Το παράθυρο αυτό εμφανίζεται είτε με το πάτημα του κουμπιού «new» είτε πληκτρολογώντας την εντολή «edit» στο παράθυρο εντολών. Αφού δημιουργήσουμε το αρχείο μπορούμε είτε να το σώσουμε στον τρέχων κατάλογο είτε να το εκτελέσουμε άμεσα με τη χρήση του κουμπιού «Run».



Εικόνα 26: Παράθυρο Σχημάτων (Figure window)

Στο Παράθυρο Σχημάτων σχεδιάζουμε γραφικές παραστάσεις με τη χρήση κατάλληλων σχεδιαστικών εντολών οι οποίες γράφονται είτε στη γραμμή εντολών είτε σε προγράμματα MATLAB.

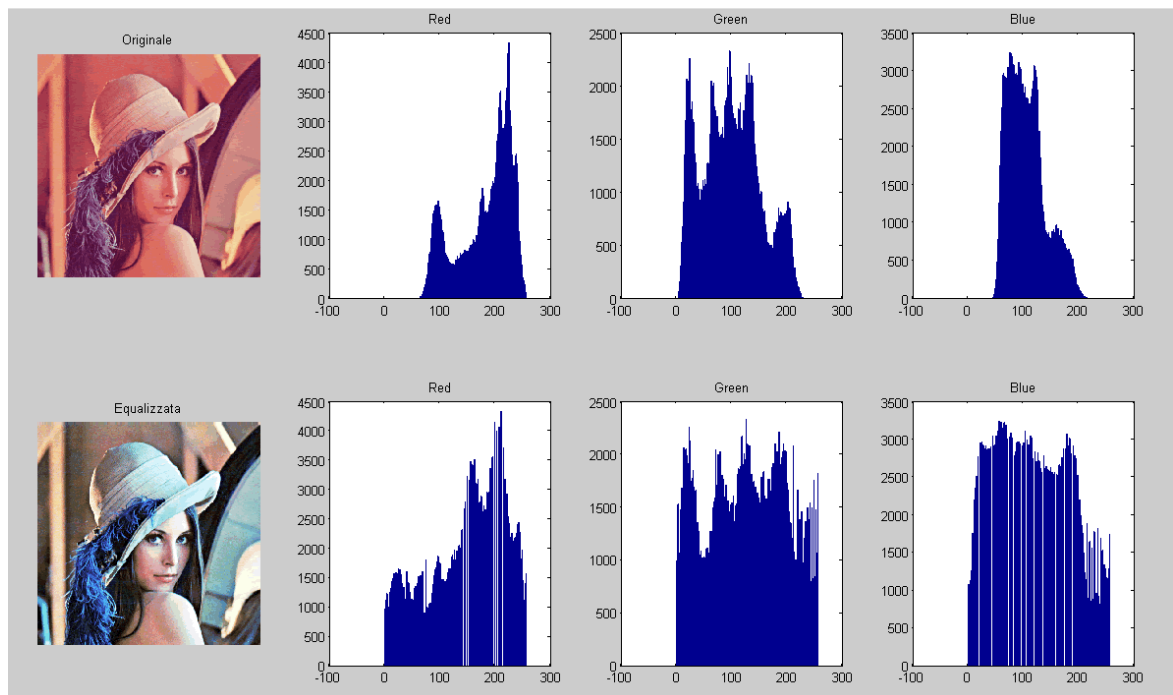


Εικόνα 27: Παράθυρο Βοήθειας (Help Window)

Το παράθυρο αυτό είναι ένα πολύ χρήσιμο εργαλείο τόσο για αρχάριους όσο και για έμπειρους χρήστες καθώς παρέχει όλες τις απαραίτητες πληροφορίες για οποιαδήποτε συνάρτηση. Μπορούμε να το χρησιμοποιήσουμε γράφοντας την εντολή «help» και το όνομα της συνάρτησης για την οποία θα θέλαμε περισσότερες πληροφορίες.

Αρχεία προγραμμάτων και συναρτήσεων

Το Matlab μας δίνει τη δυνατότητα να εργαστούμε με δύο τρόπους. Μπορούμε να χρησιμοποιήσουμε μεμονωμένες εντολές στο παράθυρο εντολών και να πάρουμε άμεσα αποτελέσματα είτε μέσω προγραμμάτων τα οποία γράφονται από το χρήστη στο παράθυρο εισαγωγής κειμένου. Τα προγράμματα αυτά αποθηκεύονται με την κατάληξη .m σαν αρχεία απλού κειμένου και ονομάζονται m-files. Τα ονόματα των m-file δεν επιτρέπεται να περιλαμβάνουν τελείες και συγκεκριμένους ειδικούς χαρακτήρες. Το πρόγραμμα δεν είναι απαραίτητο να περάσει το στάδιο της μεταγλώττισης(compiling), όπως γίνεται σε άλλες γλώσσες προγραμματισμού αλλά γίνεται αυτόματα πριν το στάδιο της εκτέλεσης. Υπάρχουν δυο ειδών τύποι m-files. Τα αρχεία προγραμμάτων(scripts) και τα αρχεία συναρτήσεων(functions). Τα αρχεία προγραμμάτων είναι αυτόνομα, δεν δέχονται ορίσματα εισόδου και δεν επιστρέφουν μεταβλητές εξόδου. Χρησιμοποιούν τις υπάρχουσες μεταβλητές που βρίσκονται στο χώρο εργασίας ή δημιουργούν καινούργιες. Τα αρχεία συναρτήσεων από την άλλη, δέχονται υποχρεωτικά ορίσματα εισόδου και μπορούν να επιστρέψουν μεταβλητές εξόδου. Το όνομα μιας συνάρτησης μαζί με τα ορίσματα εισόδου και τις μεταβλητές εξόδου γράφονται στην πρώτη γραμμή του αντίστοιχου m-file συνοδευόμενα από τη λέξη-κλειδί function, ενώ στην τελευταία γραμμή του αρχείου, γράφεται η λέξη-κλειδί end. Επίσης, το m-file που περιέχει μια συνάρτηση πρέπει να έχει το ίδιο όνομα με αυτή. Τέλος να αναφέρουμε ότι μια συνάρτηση χρησιμοποιεί το δικό της χώρο μεταβλητών, ο οποίος δεν είναι προσπελάσιμος μετά την εκτέλεσή της. Οι συναρτήσεις μπορούν να κληθούν μέσα από scripts ή από άλλες συναρτήσεις για διάφορες τιμές των ορισμάτων εισόδου τους.

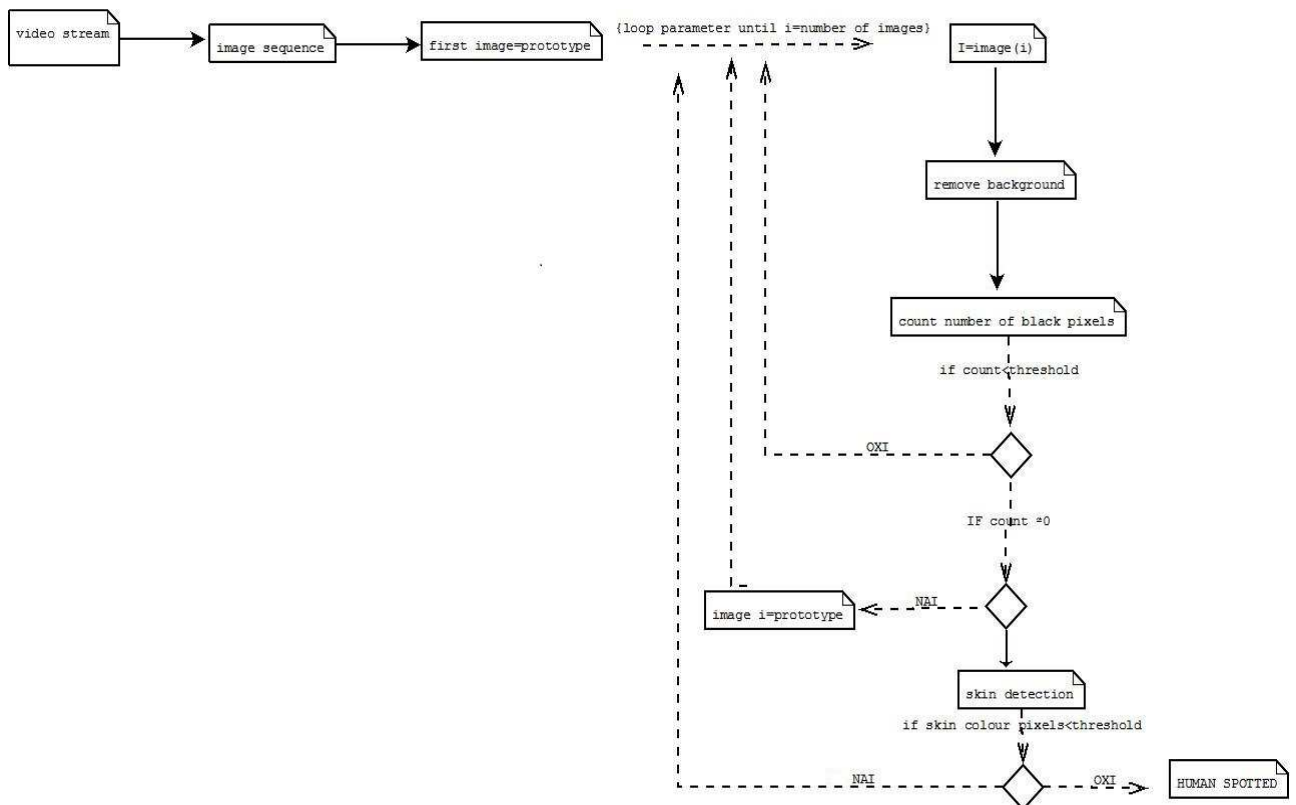


Εικόνα 28: Παράδειγμα χρήσης του Matlab στην επεξεργασία εικόνας

5.2 Εισαγωγή

Θα ξεκινήσουμε περιγράφοντας τον τρόπο με τον οποίο λειτουργεί ο αλγόριθμος και θα επισημάνουμε τις προδιαγραφές βάσει των οποίων είναι λειτουργικός, τις αδυναμίες του και τα δυνατά του σημεία. Θα εξηγήσουμε πως επιτυγχάνεται η ανίχνευση ανθρώπινης παρουσίας στο χώρο και θα δούμε διάφορα παραδείγματα του αλγόριθμου αυτού σε διάφορες σκηνές.

Ξεκινώντας, να επισημάνουμε ότι το πρόγραμμα λειτουργεί σε πλάνα τραβηγμένα από ακίνητη κάμερα, τόσο σε εσωτερικό όσο και σε εξωτερικό χώρο. Επιπλέον, είναι αρκετά ελαστικό με αλλαγές στον φωτισμό αλλά και στην εισαγωγή νέων αντικειμένων στη σκηνή χωρίς να επηρεάζεται η αποδοτικότητα του.



Εικόνα 29: Γραφική Αναπαράσταση Αλγορίθμου

Αρχικά, ο χρήστης επιλέγει το βίντεο στο οποίο θέλει να εφαρμόσει τον αλγόριθμο για να ελέγξει για παρουσία ανθρώπου. Το βίντεο αυτό αναπαράγεται και χωρίζεται σε στιγμιότυπα για μετέπειτα επεξεργασία. Απαραίτητη προϋπόθεση είναι να μην υπάρχει στη σκηνή ανθρώπινη παρουσία στο πρώτο frame. Αφού δημιουργηθούν τα στιγμιότυπα ξεκινάει ο αλγόριθμος που ελέγχει τις εικόνες για τον εντοπισμό ανθρώπου. Ορίζει το πρώτο frame του βίντεο σαν πρωτότυπο και ξεκινάει την διαδικασία της αφαίρεσης background διαβάζοντας τις εικόνες και συγκρίνοντας τις με την πρώτη. Στη συνέχεια ελέγχει το ποσοστό της εικόνας που έχει αλλάξει, και αν είναι μεγαλύτερο από ένα συγκεκριμένο κατώφλι που ορίζουμε εμείς προχωράει στο δεύτερο μέρος, δηλαδή τη διαδικασία για εντοπισμό pixel στο χρώμα του δέρματος. Στην περίπτωση που το ποσοστό που άλλαξε είναι αρκετά μικρό θεωρούμε ότι δεν υπάρχει άνθρωπος στη σκηνή στο συγκεκριμένο frame και προχωράει στο επόμενο. Προκειμένου να αποφύγουμε λάθη που προκύπτουν από την παρουσία χρώματος παρόμοιου με αυτό του ανθρώπινου δέρματος, μελετάμε την εικόνα που έχει προκύψει από την αφαίρεση του background. Με αυτό τον τρόπο έχουμε μικρότερη

πιθανότητα για εσφαλμένη ανίχνευση, αφού εργαζόμαστε μόνο πάνω στο κομμάτι της εικόνας που έχουμε εντοπίσει διαφορά σε σχέση με την αρχική μας. Αν ο αλγόριθμος εντοπίσει pixel με χρώμα κοντινό σε αυτό του δέρματος, θεωρεί ότι η αλλαγή στην εικόνα προέκυψε από ανθρώπινη παρουσία και τερματίζει τη λειτουργία του επιστρέφοντας στην εικόνα που εντόπισε τον άνθρωπο, καθώς και στην εικόνα που προέκυψε από την αφαίρεση του φόντου. Στην περίπτωση που δεν εντοπίσει pixel στο χρώμα του δέρματος, υποθέτουμε ότι υπήρξε μια αλλαγή στην εικόνα μας η οποία δεν ήταν αποτέλεσμα ανθρώπινης παρουσίας και συνεχίζει την εκτέλεση του προγράμματος ελέγχοντας το επόμενο frame. Σε αυτό το σημείο, να κάνουμε μια περιγραφή των δυο βασικών προβλημάτων που προκύπτουν. Το ένα είναι το πρόβλημα της αλλαγής φωτισμού και το δεύτερο της μετακίνησης σημείων του background.

Στην πρώτη περίπτωση, όταν αλλάζει ο φωτισμός απότομα, το πρόγραμμα θα εντοπίσει μεγάλη αλλαγή στην εικόνα μας και θα ξεκινήσει τη διαδικασία για ανίχνευση δέρματος. Η πιθανότητα να βρει σε όλη την εικόνα pixel στο χρώμα του δέρματος είναι αρκετά μεγάλη λόγω των αντικειμένων που θα βρίσκονται στο χώρο και θα έχουν παρόμοιο χρώμα με αυτό του δέρματος, για παράδειγμα ξύλινα αντικείμενα. Για να αποφύγουμε αυτό το πρόβλημα εισάγουμε μια νέα συνθήκη στον κώδικα έτσι ώστε όταν ανιχνεύσει αλλαγή σε όλα τα pixel της εικόνας να θεωρεί ότι η αλλαγή αυτή προήλθε από αλλαγή του φωτισμού. Σε αυτή την περίπτωση θεωρεί ως πρωτότυπη την εικόνα, η οποία προέκυψε από την αλλαγή φωτισμού και συνεχίζει να ελέγχει τις επόμενες βάσει αυτής.

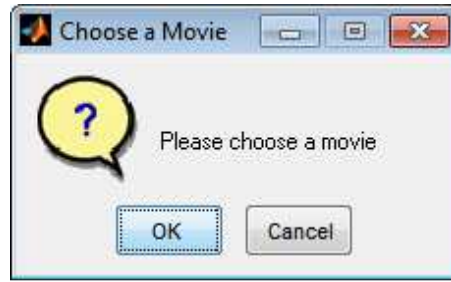
Το δεύτερο πρόβλημα αφορά σκηνές που δεν είναι ελεγχόμενες. Όπως είναι οι σκηνές εξωτερικού χώρου. Σε ένα τέτοιο περιβάλλον υπάρχουν αρκετοί αντιπερισπασμοί, όπως είναι τα φύλλα των δέντρων, η κίνηση οχημάτων και διάφορα ζώα. Για να κάνουμε το σύστημα μας πιο ελαστικό σε ένα τέτοιο περιβάλλον προσπαθήσαμε να κάνουμε τον αλγόριθμο της αφαίρεσης background να αλλάζει πιο ομαλά από ότι θα άλλαζε με την αφαίρεση της μιας εικόνας από την άλλη. Η συνάρτηση που προκύπτει μετά από μορφοποίηση είναι η εξής:

$$\text{Background} = (1-\alpha) * I + \alpha * I_b;$$

Στην συνάρτηση αυτή η εικόνα με τίτλο background, δημιουργείται από ένα ποσοστό του αρχικού frame και ένα ποσοστό του frame το οποίο εξετάζουμε για παρουσία ανθρώπου. Με αυτό τον τρόπο σκοπεύουμε να κάνουμε τον αλγόριθμό μας πιο ελαστικό σε μικρές αλλαγές, που προκύπτουν από κινήσεις που δεν παρουσιάζουν ενδιαφέρον και εντάσσονται στα στοιχεία του background. Μετά από αυτή την περιγραφή θα δούμε κατά πόσο είναι λειτουργικός ο αλγόριθμος σε πραγματικά περιβάλλοντα και θα εντοπίσουμε τυχόν αδυναμίες του αλγόριθμου.

5.3 Εφαρμογή

Θα ξεκινήσουμε την παρουσίαση της εφαρμογής με την δοκιμή του αλγόριθμου στο πιο απλό περιβάλλον. Σε αυτή την περίπτωση θα μελετήσουμε πως λειτουργεί σε ένα ελεγχόμενο εσωτερικό χώρο χωρίς κανένα περισπασμό. Σε αυτό το πλάνο θα έχουμε μια σταθερή κάμερα με σταθερό και ελεγχόμενο φωτισμό και το αντικείμενο ενδιαφέροντος θα περάσει μπροστά από το πλάνο προκειμένου να δούμε αν θα καταφέρει να το αναγνωρίσει. Όταν πατήσουμε το πλήκτρο για να εκτελεστεί ο κώδικας εμφανίζεται ένα παράθυρο για να μπορέσει ο χρήστης να διαλέξει το βίντεο που επιθυμεί.



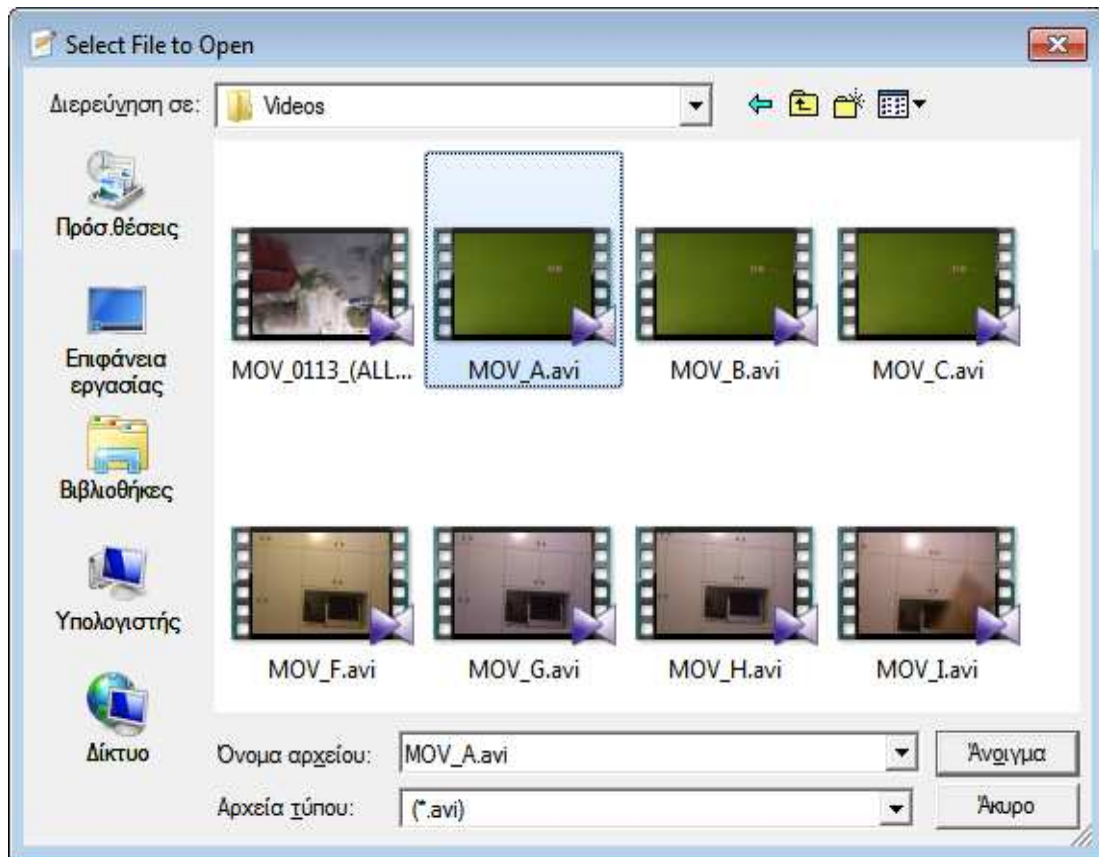
Εικόνα 30: Παράθυρο επιλογής βίντεο

Τα βίντεο πρέπει να είναι σε μορφή που υποστηρίζεται από το Matlab. Ακολουθεί ένας πίνακας που δείχνει τα format που υποστηρίζονται από το Matlab

All Platforms	AVI (.avi), Motion JPEG 2000 (.mj2)
All Windows®	MPEG-1 (.mpg), Windows Media® Video (.wmv, .asf, .asx), and any format supported by Microsoft® DirectShow®.
Windows 7	MPEG-4, including H.264 encoded video (.mp4, .m4v), Apple QuickTime Movie (.mov), and any format supported by Microsoft Media Foundation.
Macintosh	MPEG-1 (.mpg), MPEG-4, including H.264 encoded video (.mp4, .m4v), Apple QuickTime Movie (.mov), and any format supported by QuickTime as listed on http://support.apple.com/kb/HT3775 .
Linux®	Any format supported by your installed plug-ins for GStreamer 0.10 or above, as listed on http://gstreamer.freedesktop.org/documentation/plugins.html , including Ogg Theora (.ogg).

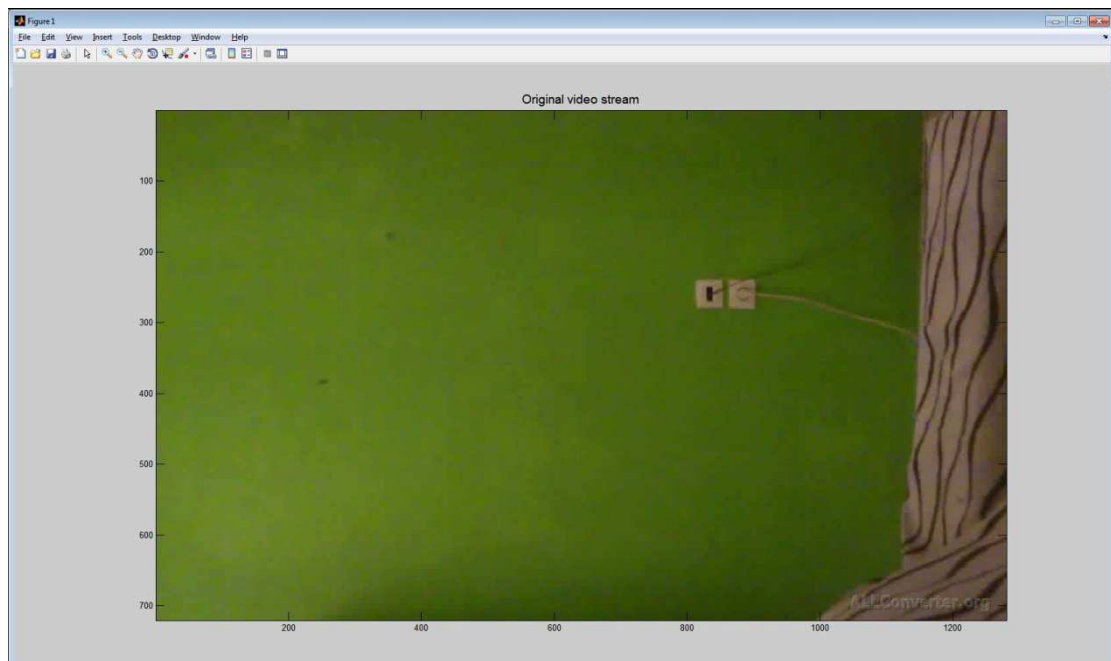
Πίνακας υποστηριζόμενων μορφών βίντεο για το Matlab

Στη συνέχεια ο χρήστης διαλέγει το βίντεο που επιθυμεί, το οποίο εισάγεται στο Matlab για να αναλυθεί και να εξεταστεί για τυχόν παρουσία ανθρώπου.

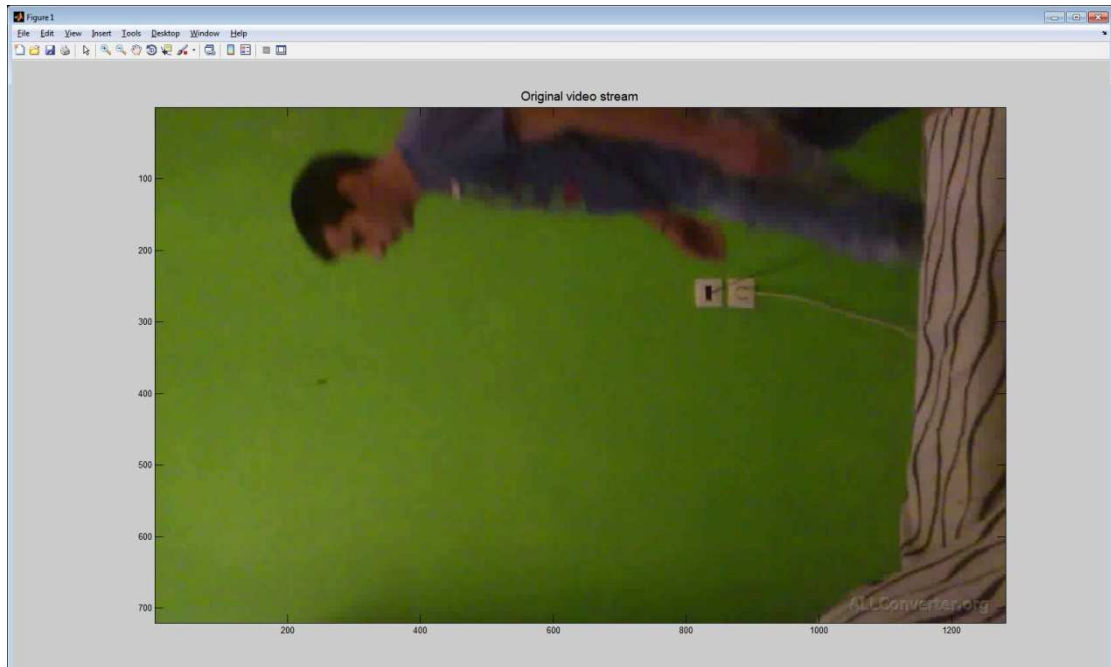


Εικόνα 31: Επιλογή αρχείου βίντεο προς επεξεργασία

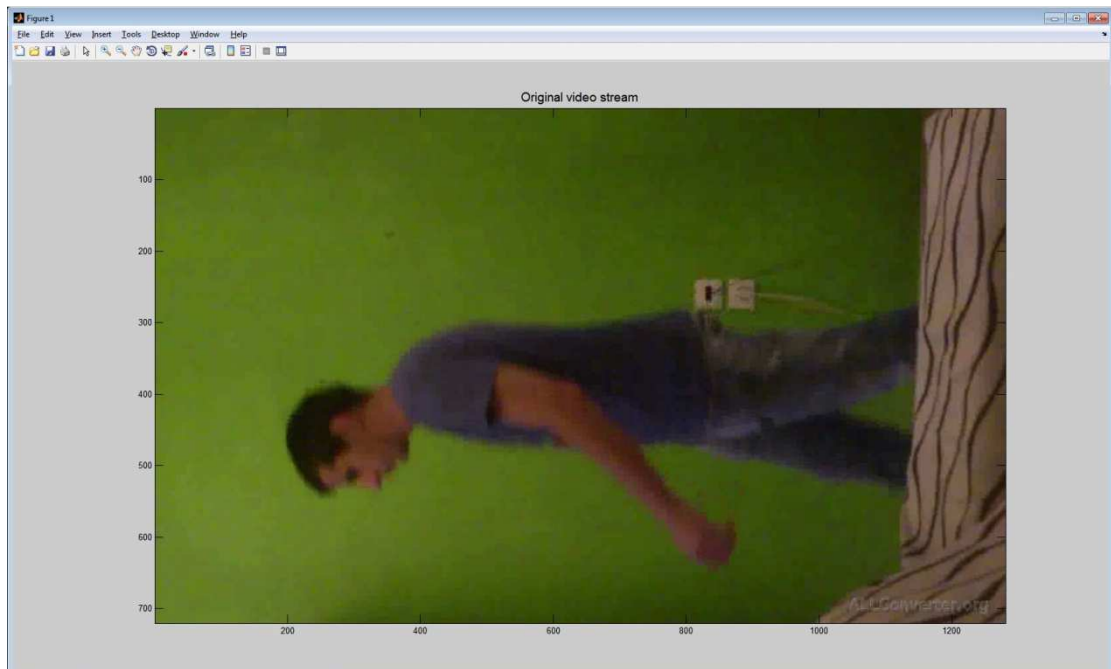
Αφού επιλέξουμε το αρχείο που επιθυμούμε ξεκινάει η αναπαραγωγή του βίντεο, ενώ ταυτόχρονα αποθηκεύεται η ακολουθία εικόνων.



Εικόνα 32: Στιγμιότυπο 1 από την αναπαραγωγή βίντεο1

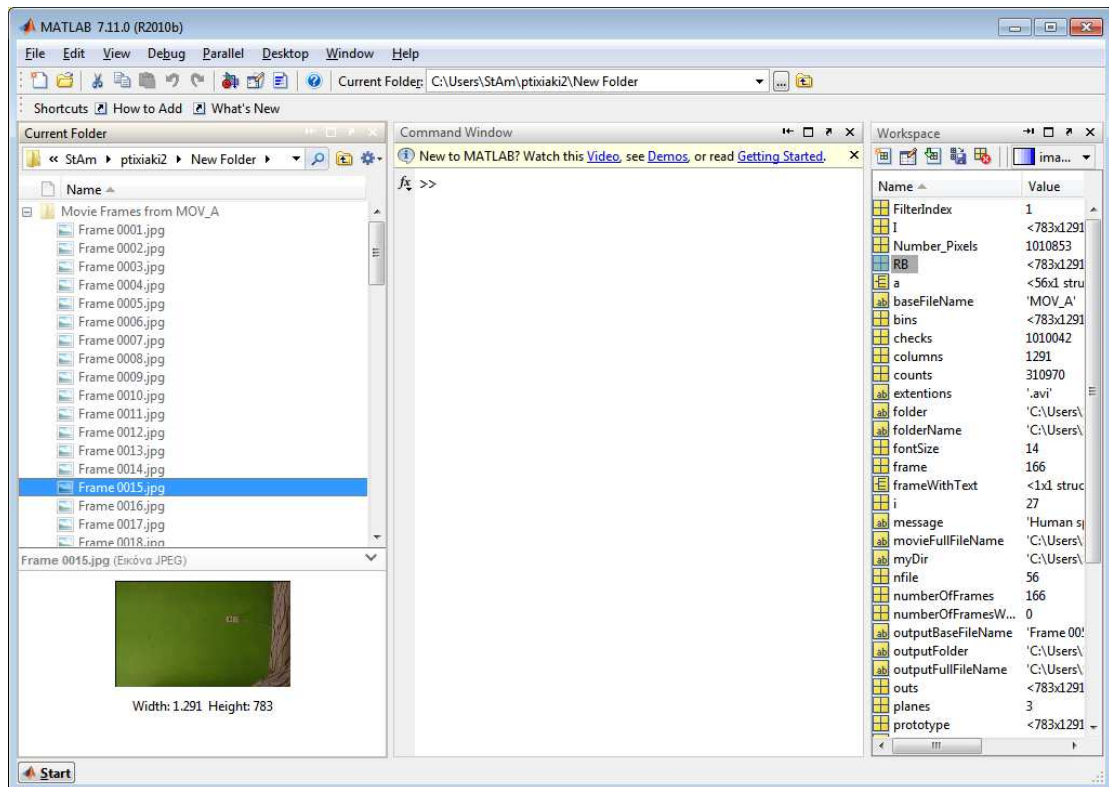


Εικόνα 33: Στιγμιότυπο 2 από την αναπαραγωγή βίντεο 1



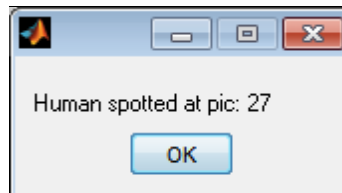
Εικόνα 34: Στιγμιότυπο 3 από την αναπαραγωγή βίντεο 1

Βλέπουμε στην παρακάτω εικόνα ότι στο φάκελο εργασίας μας στο Matlab έχει δημιουργηθεί ένας καινούργιος φάκελος με το όνομα του βίντεο και μέσα σε αυτόν βρίσκονται όλα τα στιγμιότυπα από το βίντεο.

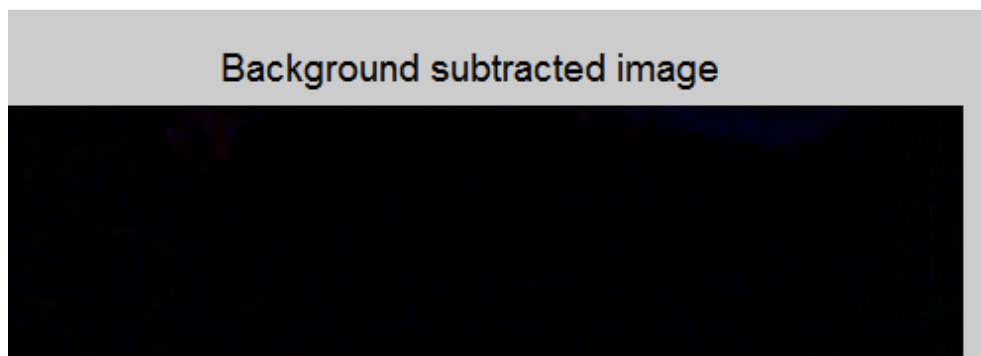


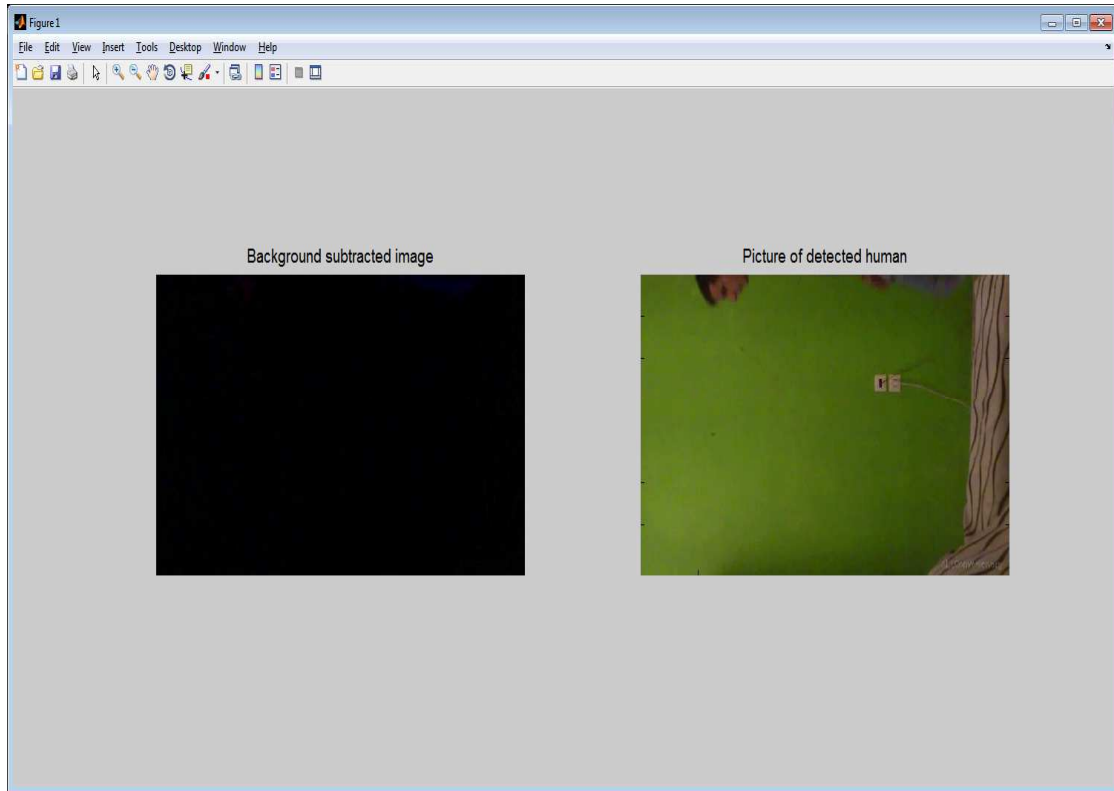
Εικόνα 35: Στιγμιότυπο από το περιβάλλον εργασίας του Matlab

Μόλις τελειώσει η αναπαραγωγή του βίντεο ξεκινάει η ανίχνευση για ανθρώπινη παρουσία, με τον τρόπο που περιγράψαμε παραπάνω. Αν καταφέρει να εντοπίσει ανθρώπινη παρουσία εμφανίζει ένα μήνυμα με τον αριθμό της εικόνας που εντόπισε τον άνθρωπο, καθώς και την αρχική εικόνα μαζί με αυτήν που προέκυψε από την αφαίρεση του background.

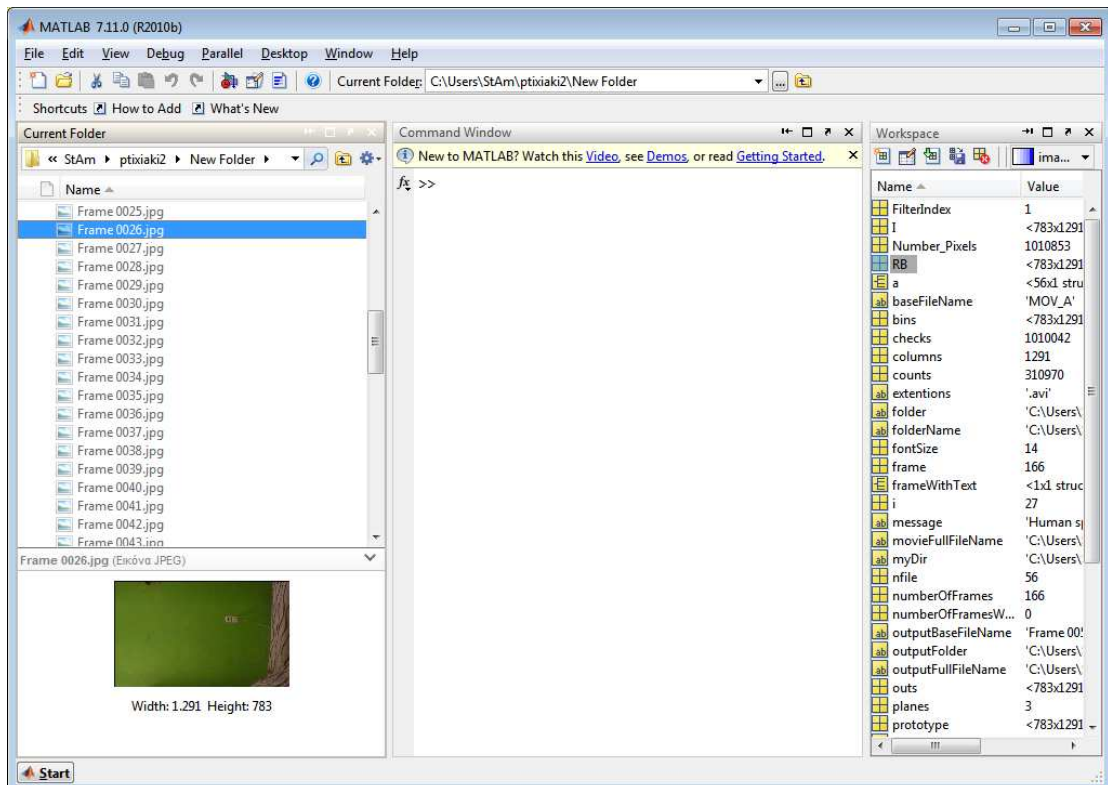


Εικόνα 36: Μήνυμα ανίχνευσης ανθρώπου στο βίντεο 1





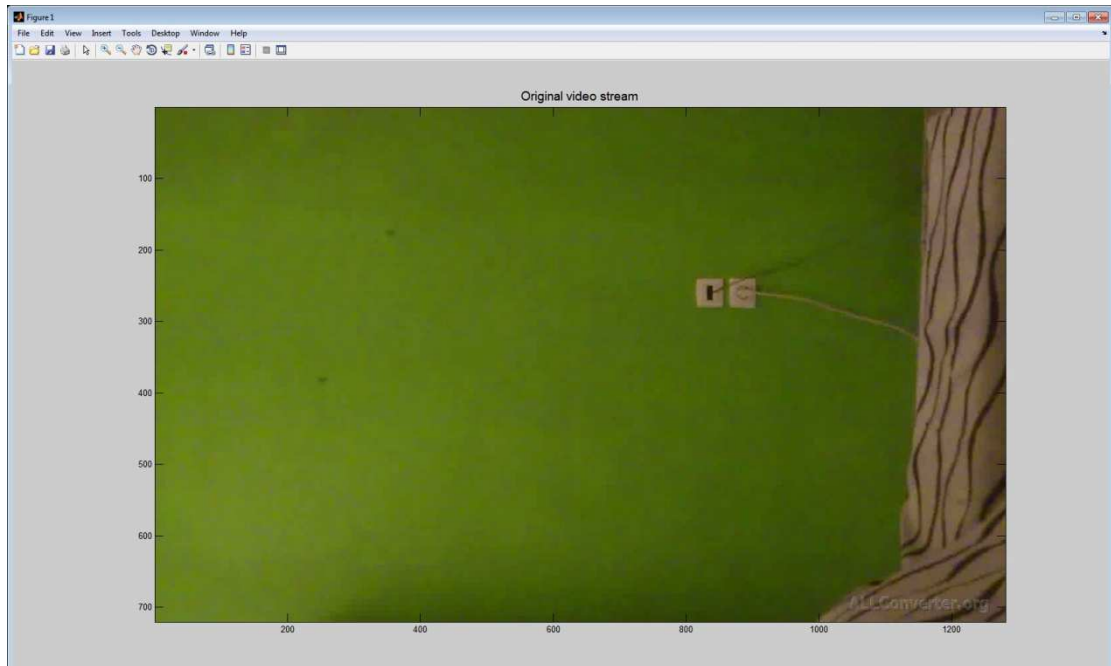
Εικόνα 37: Η εικόνα που εντοπίστηκε ανθρώπινη παρουσία με και χωρίς αφαίρεση background στο βίντεο 1



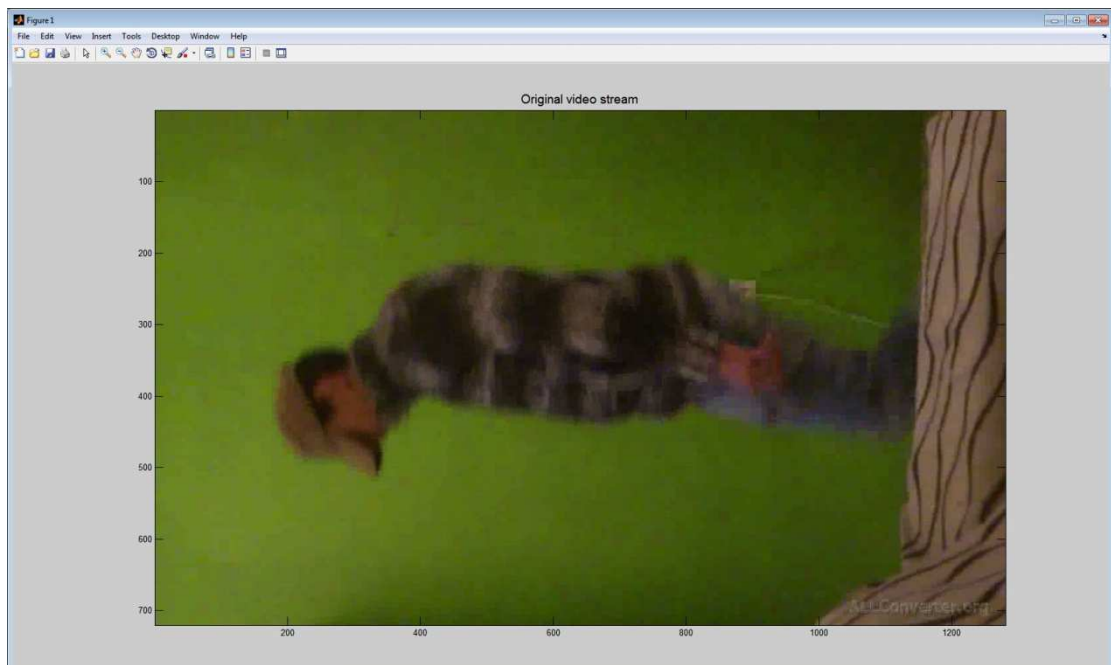
Εικόνα 38: Στιγμιότυπο από το περιβάλλον εργασίας του Matlab μετά την εκτέλεση της εφαρμογής για το βίντεο 1

Βλέπουμε στην παραπάνω εικόνα ότι το frame 27 είναι πράγματι η πρώτη φορά που εμφανίζεται η ανθρώπινη παρουσία. Συνεπώς, συμπεραίνουμε ότι ο αλγόριθμος είναι

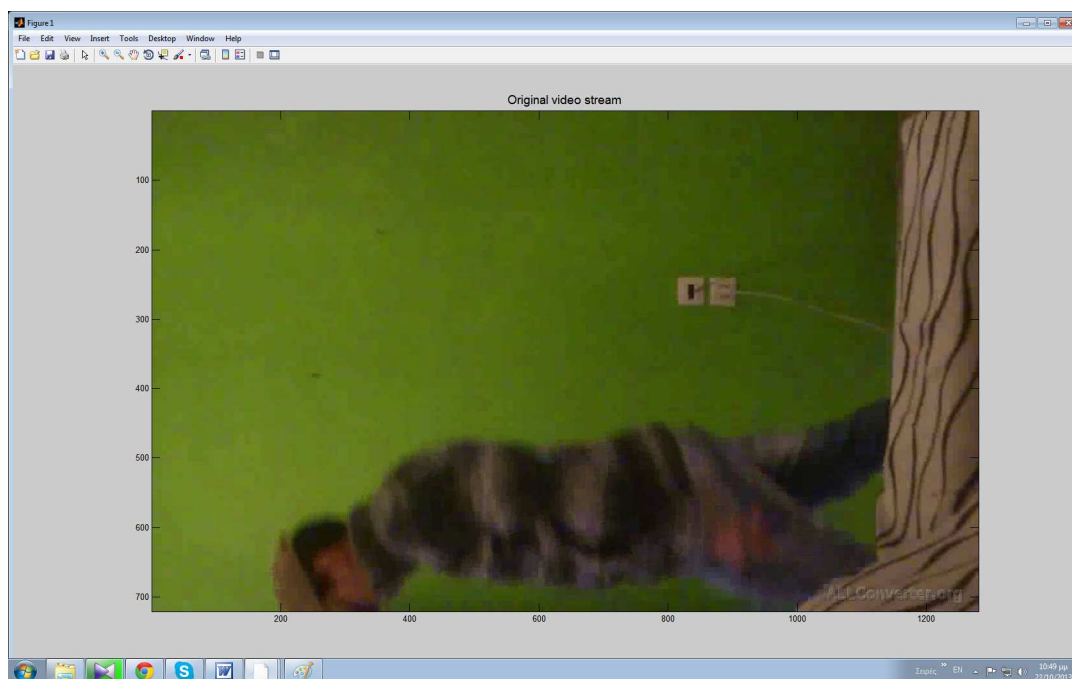
αρκετά αποδοτικός σε περιβάλλοντα πάρα πολύ απλά χωρίς περισπασμούς. Εν συνεχεία, θα δοκιμάσουμε το αν μπορεί να ανιχνεύσει ανθρώπινη παρουσία όταν είναι λιγότερα τα pixels του δέρματος. Στο επόμενο βίντεο προσπαθούμε να μειώσουμε τις περιοχές που εμφανίζεται δέρμα.



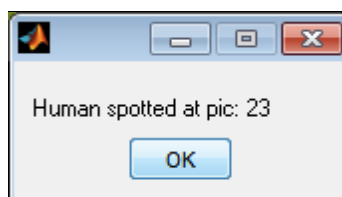
Εικόνα 39: Στιγμιότυπο 1 από την αναπαραγωγή βίντεο 2



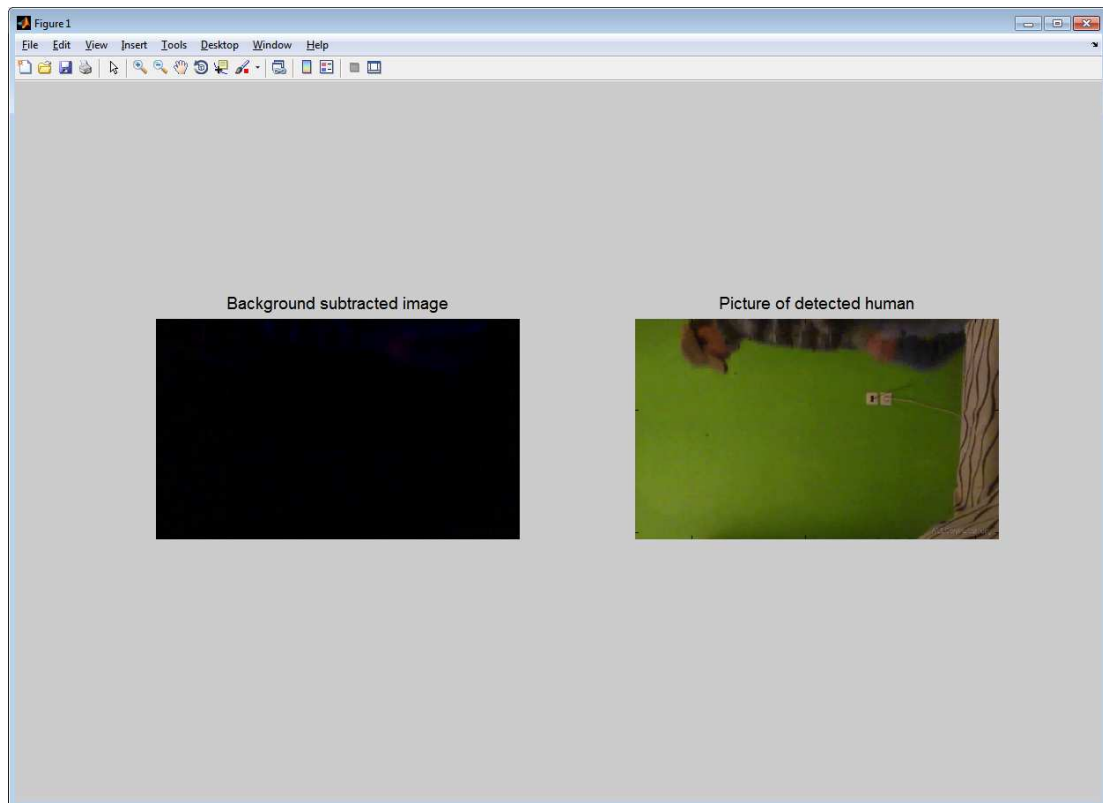
Εικόνα 40: Στιγμιότυπο 2 από την αναπαραγωγή βίντεο 2



Εικόνα 41: Στιγμιότυπο 3 από την αναπαραγωγή βίντεο 2

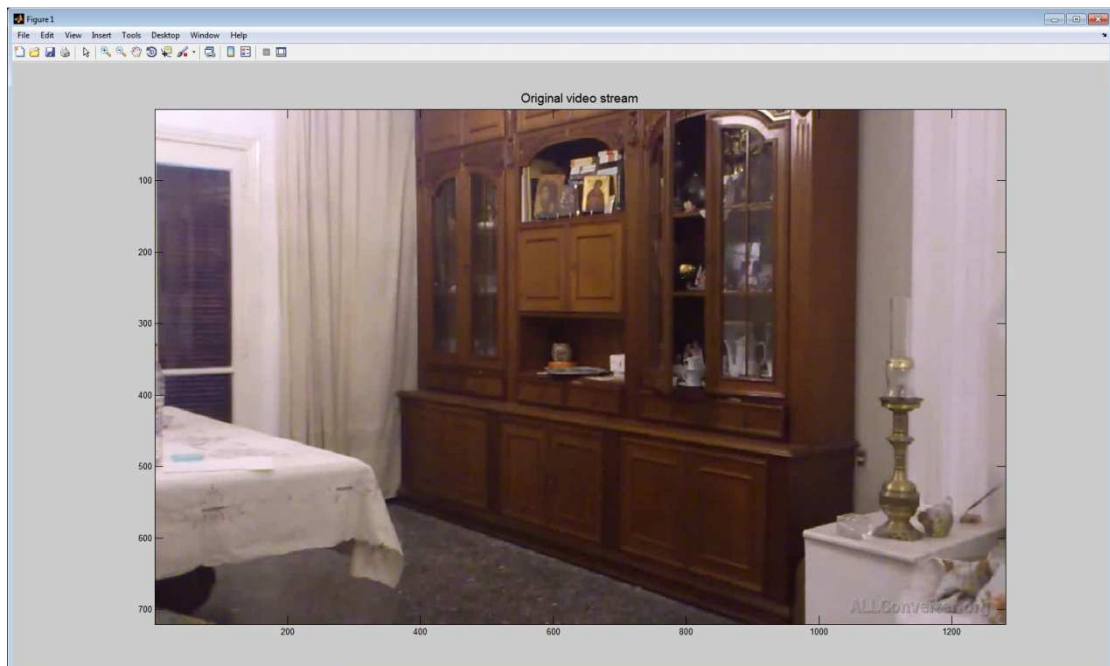


Εικόνα 42: Μήνυμα αντίγνωσης ανθρώπου στο βίντεο 2

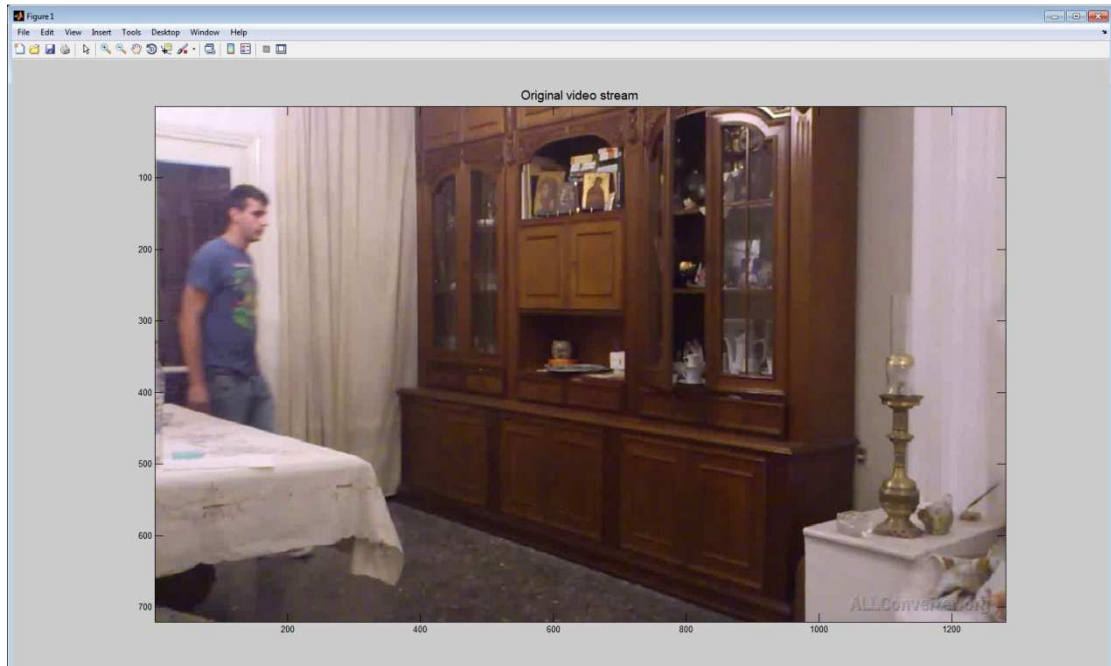


Εικόνα 43: Η εικόνα που εντοπίστηκε ανθρώπινη παρουσία στο βίντεο 2 με και χωρίς αφαίρεση background

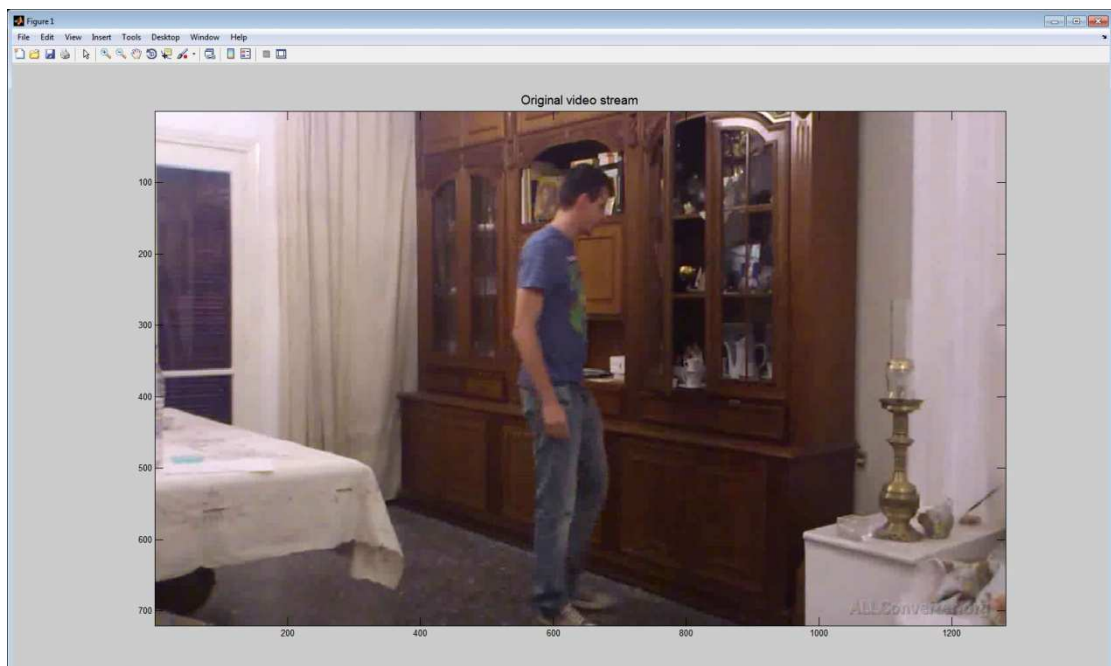
Βλέπουμε ότι και με τις περισσότερες περιοχές δέρματος καλυμμένες ο αλγόριθμος καταφέρνει με επιτυχία να ανιχνεύσει την ανθρώπινη παρουσία. Στο επόμενο στάδιο θα δοκιμάσουμε το πρόγραμμα με ένα βίντεο τραβηγμένο από ένα πιο περίπλοκο εσωτερικό χώρο, όπου θα υπάρχουν πολλά αντικείμενα στο χρώμα του δέρματος.



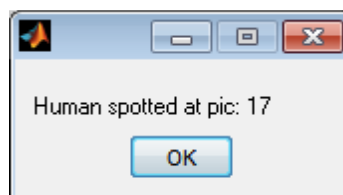
Εικόνα 44: Στιγμιότυπο 1 από την αναπαραγωγή βίντεο 3



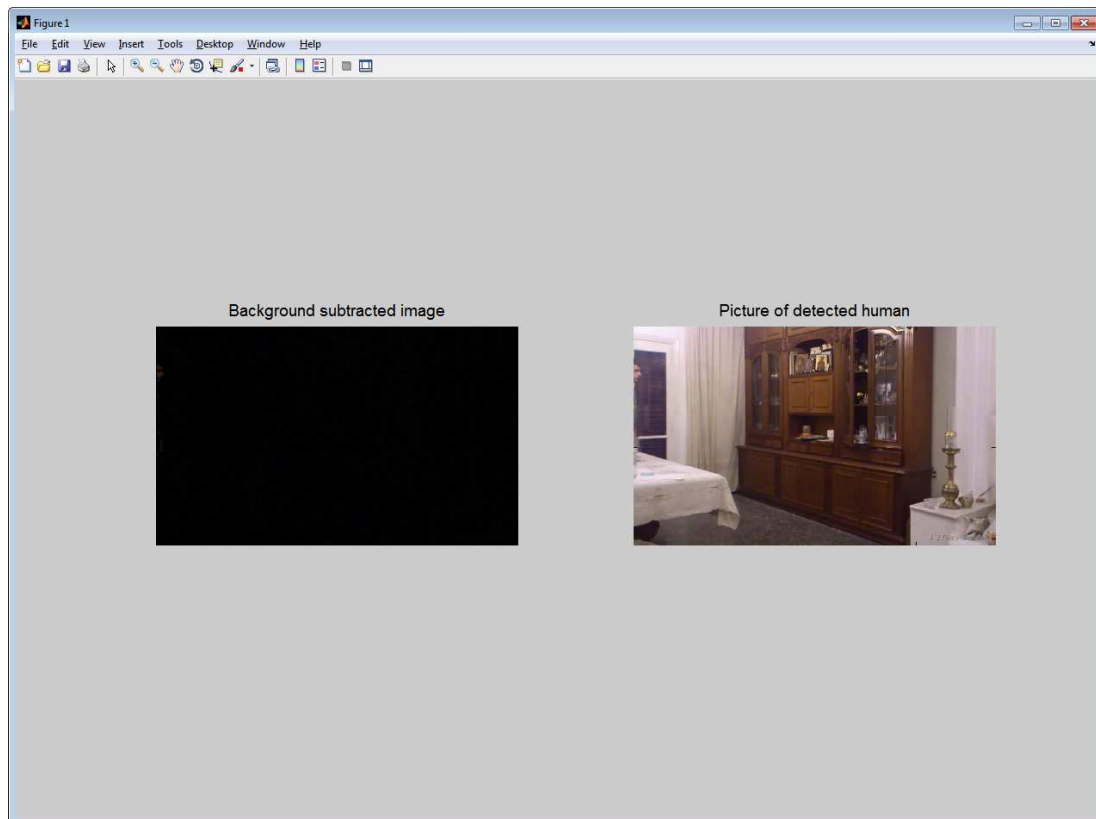
Εικόνα 45: Στιγμιότυπο 2 από την αναπαραγωγή βίντεο 3



Εικόνα 46: Στιγμιότυπο 3 από την αναπαραγωγή βίντεο 3

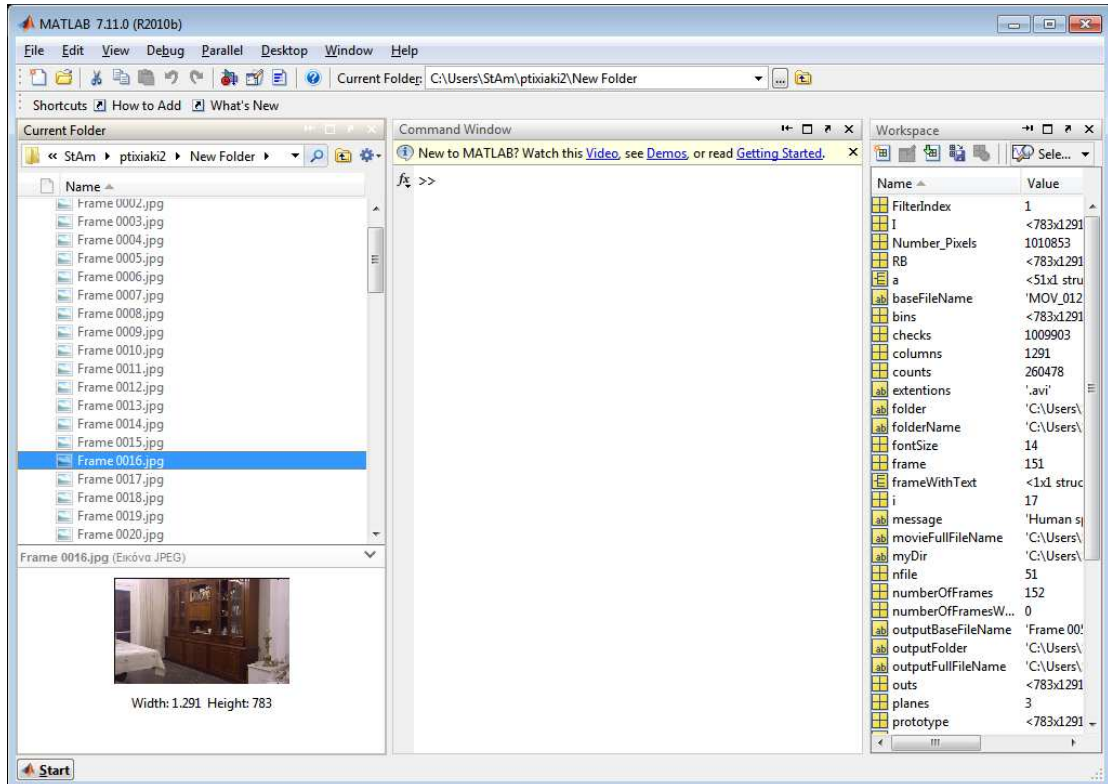


Εικόνα 47: Μήνυμα ανίχνευσης ανθρώπου στο βίντεο 3



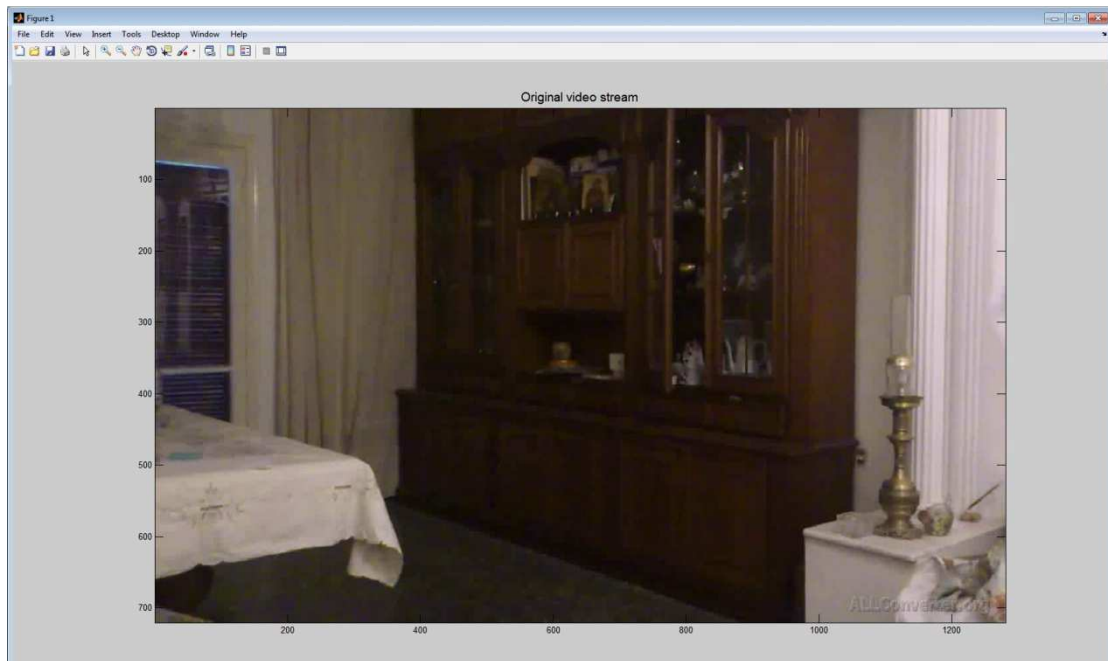
Εικόνα 48: Η εικόνα που εντοπίστηκε ανθρώπινη παρουσία στο βίντεο 3 με και χωρίς αφαίρεση background

Βλέπουμε ότι παρόλη την πολυπλοκότητα του φόντου ο αλγόριθμος με επιτυχία εντοπίζει την ανθρώπινη παρουσία, μόλις εμφανίζεται, και δεν επηρεάζεται από τις σκιάσεις και τις ανακλάσεις που προκύπτουν σταδιακά στην εικόνα.

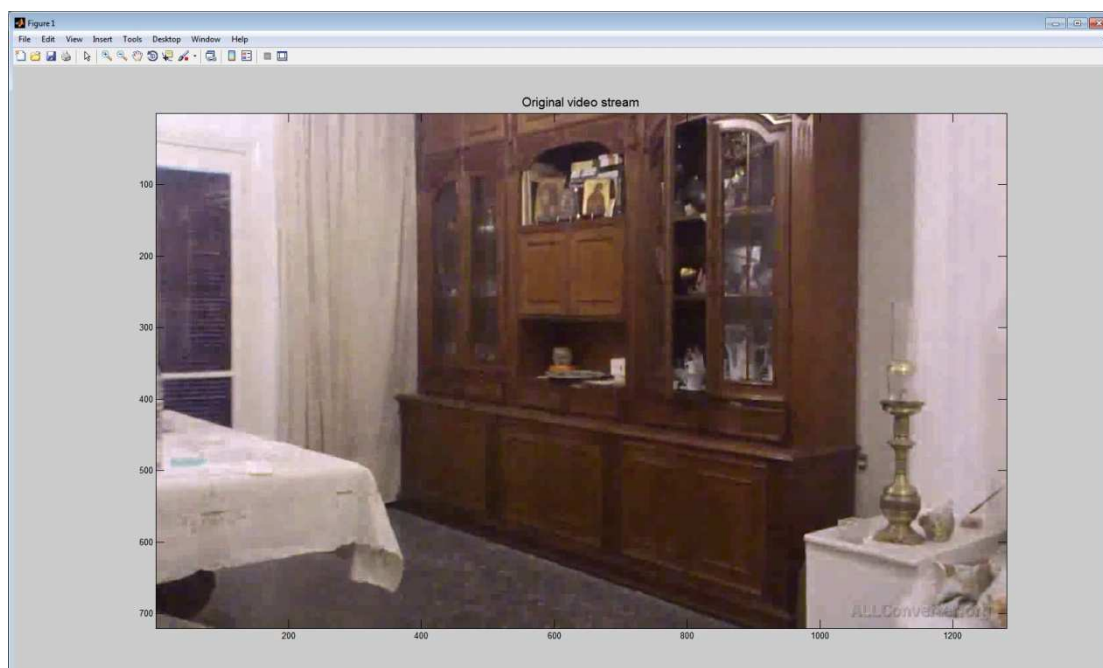


Εικόνα 49: Στιγμιότυπο από το περιβάλλον εργασίας του Matlab μετά την εκτέλεση της εφαρμογής για το βίντεο 3

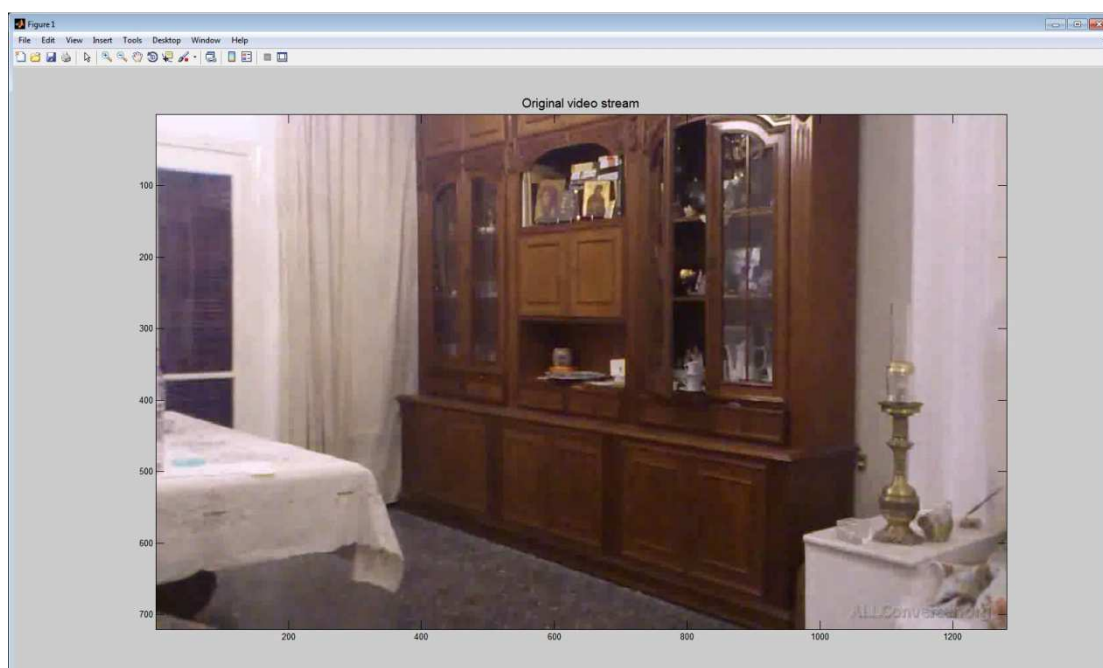
Στη συνέχεια θα δοκιμάσουμε την αποτελεσματικότητά του σε περιβάλλον με αλλαγή φωτισμού. Στο δωμάτιο θα αυξήσουμε το φωτισμό και στη συνέχεια θα εμφανιστεί ο άνθρωπος.



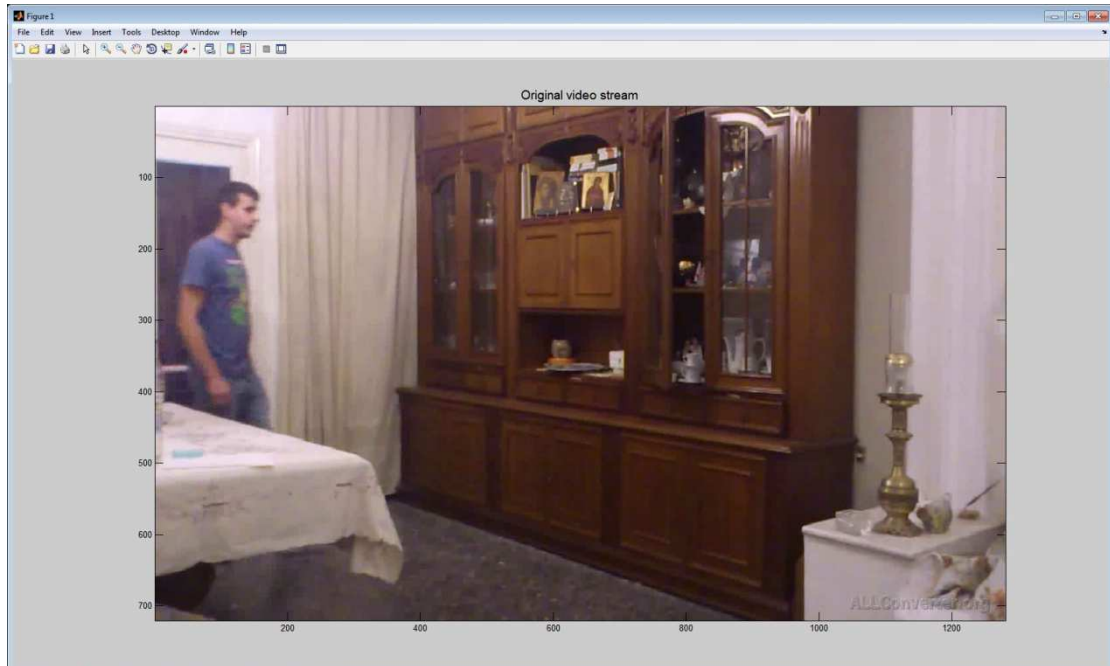
Εικόνα 50: Στιγμιότυπο 1 από την αναπαραγωγή βίντεο 4



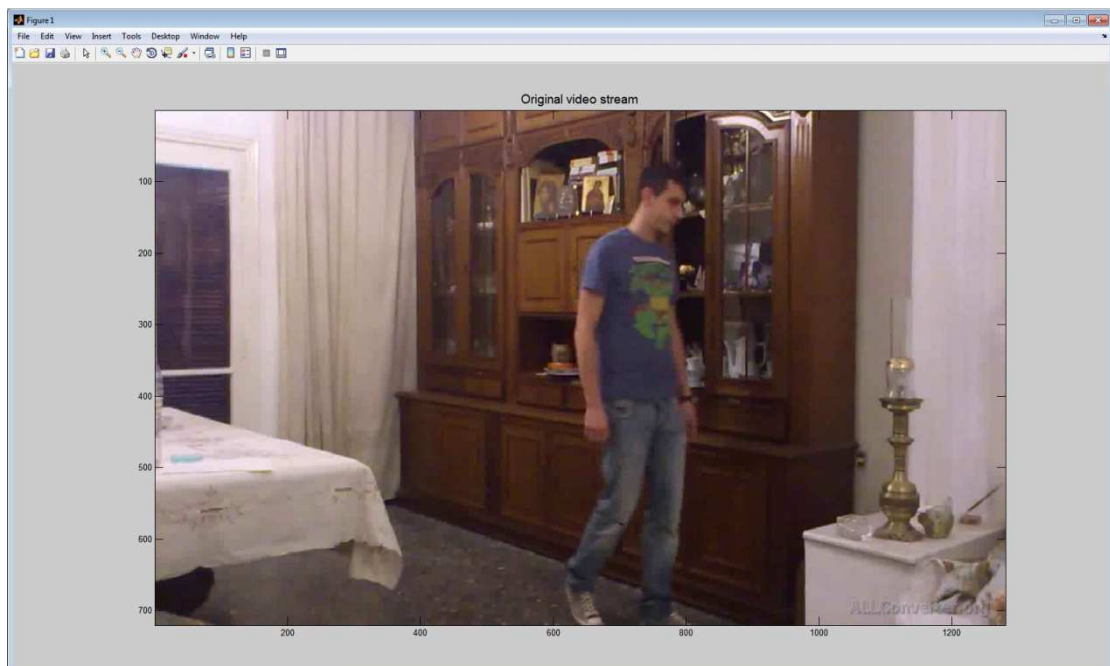
Εικόνα 51: Στιγμιότυπο 2 από την αναπαραγωγή βίντεο 4



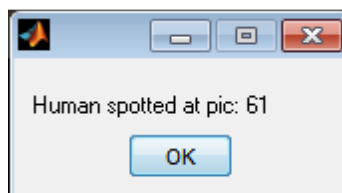
Εικόνα 52: Στιγμιότυπο 3 από την αναπαραγωγή βίντεο 4



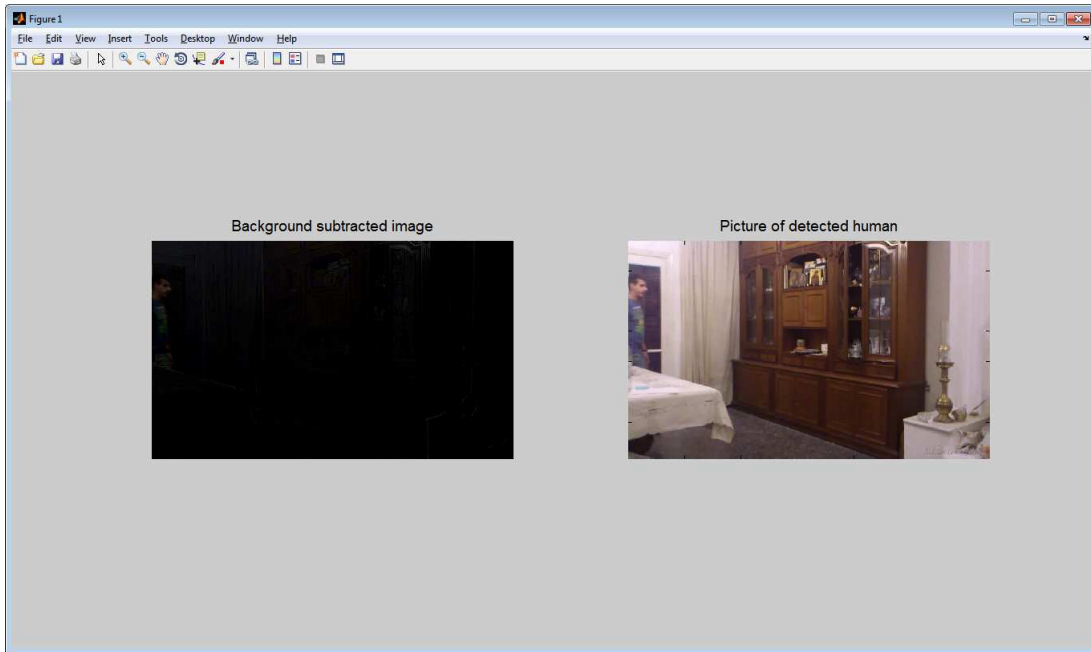
Εικόνα 53: Στιγμιότυπο 4 από την αναπαραγωγή βίντεο 4



Εικόνα 54: Στιγμιότυπο 5 από την αναπαραγωγή βίντεο 4

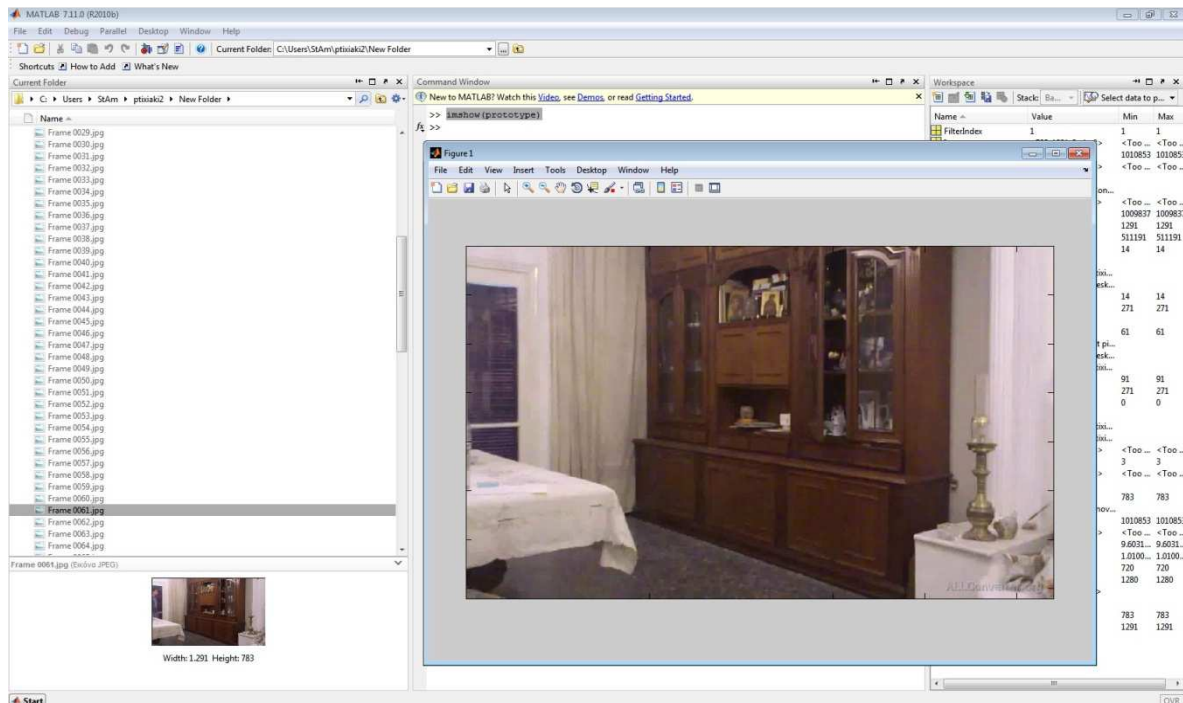


Εικόνα 55: Μήνυμα ανίχνευσης ανθρώπου στο βίντεο 4



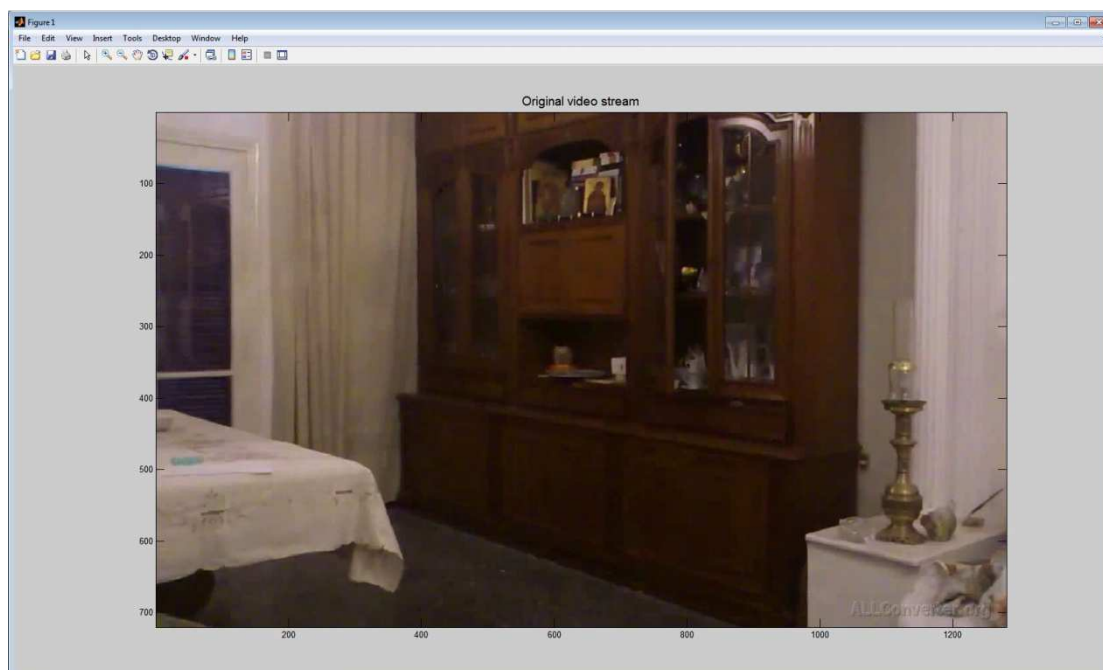
Εικόνα 56: Η εικόνα που εντοπίστηκε ανθρώπινη παρουσία στο βίντεο 4 με και χωρίς αφαίρεση background

Παρατηρούμε ότι παρόλο που αλλάζει ο φωτισμός το πρόγραμμα δεν κάνει εσφαλμένη ανίχνευση σε pixels που έχουν χρώμα παρόμοιο με αυτό του ανθρώπινου δέρματος. Αφού αλλάζει ο φωτισμός όλα τα pixels τις εικόνες αλλάζουν χρώμα και δεν ξεκινάει τον αλγόριθμο για ανίχνευση δέρματος αλλά θεωρεί ως πρωτότυπη εικόνα την καινούργια και συνεχίζει την εκτέλεση του όπως βλέπουμε και στην εικόνα που ακολουθεί.

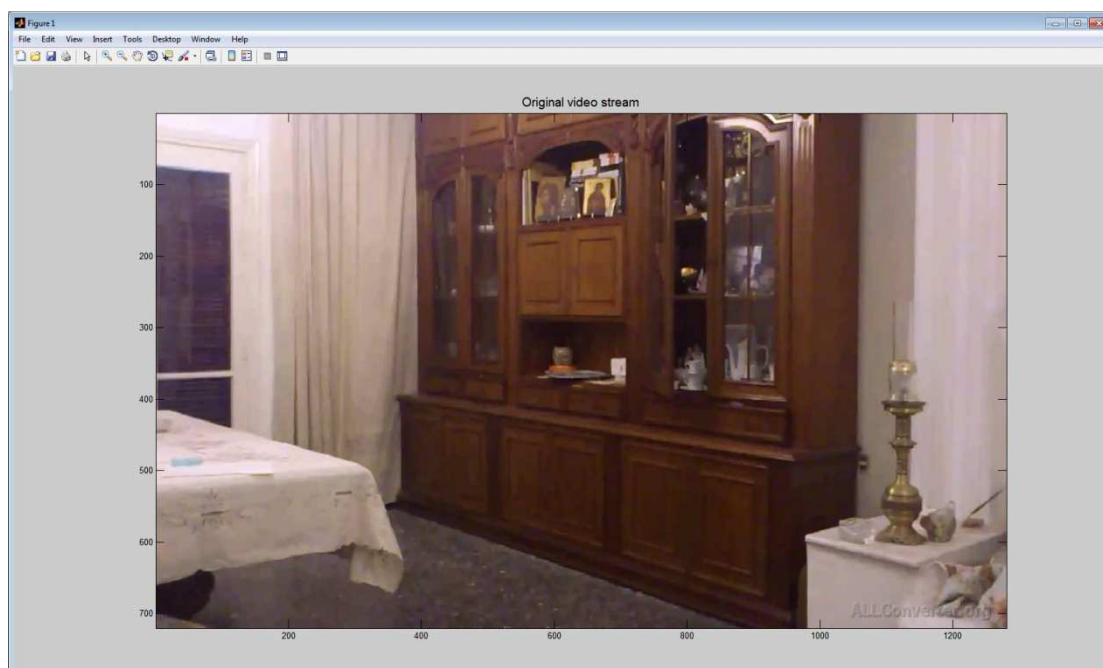


Εικόνα 57: Εναλλαγή αρχικής εικόνας όταν αλλάζει ο φωτισμός

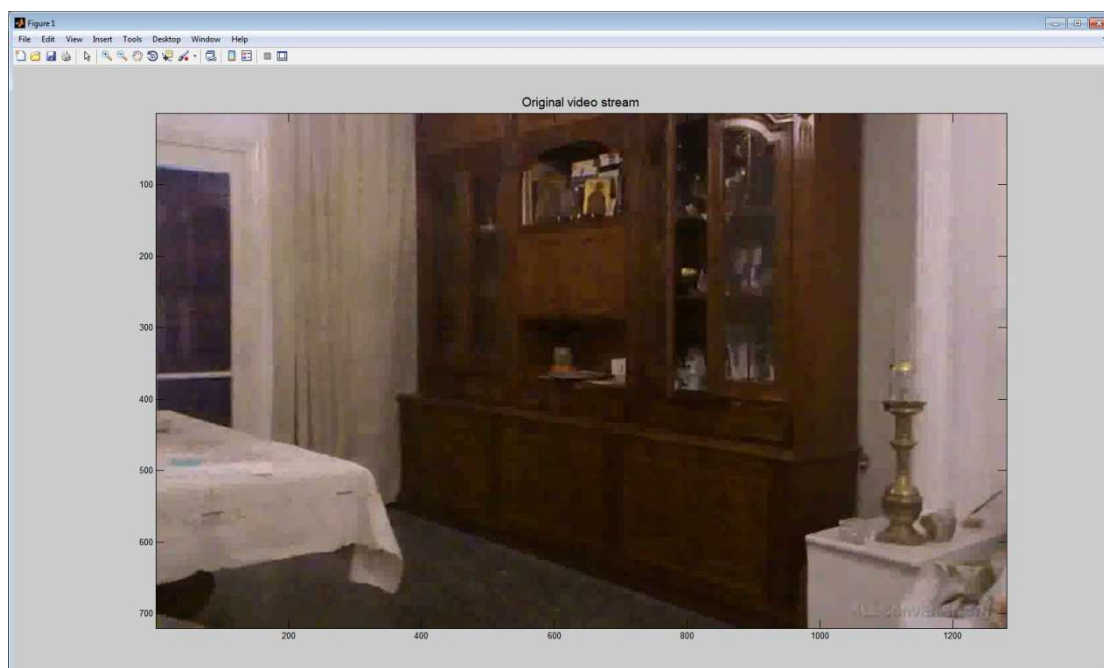
Τελειώνοντας με τις σκηνές εσωτερικού χώρου θα δοκιμάσουμε τι θα γίνει στην περίπτωση που ο φωτισμός δεν μένει σταθερός αλλά αλλάζει συνεχώς.



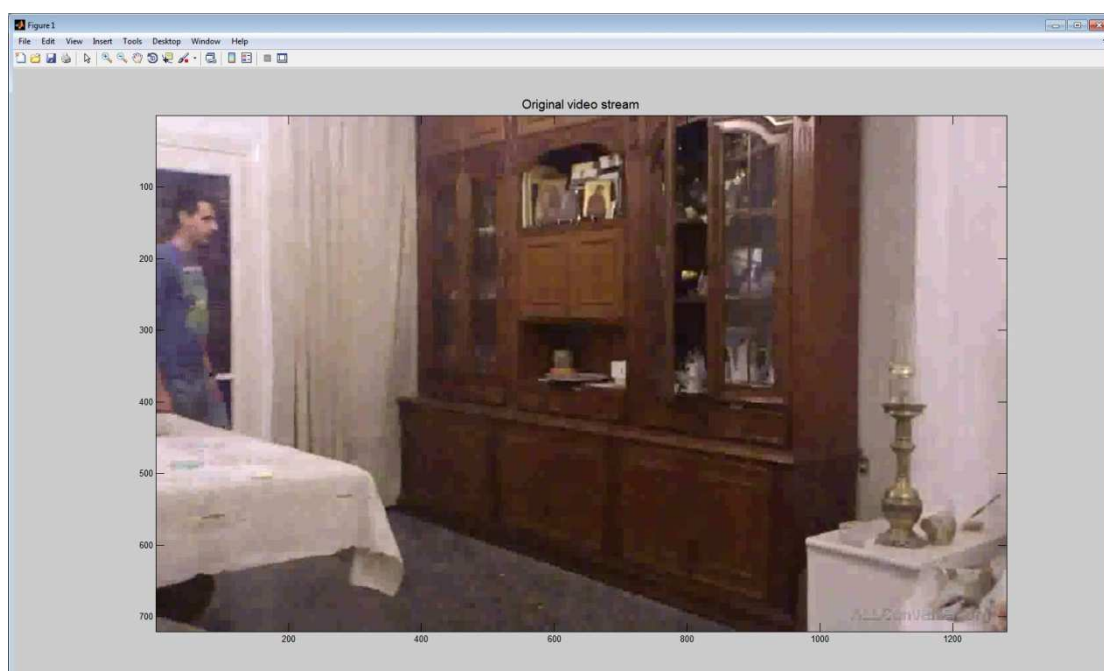
Εικόνα 58: Στιγμιότυπο 1 από την αναπαραγωγή βίντεο 5



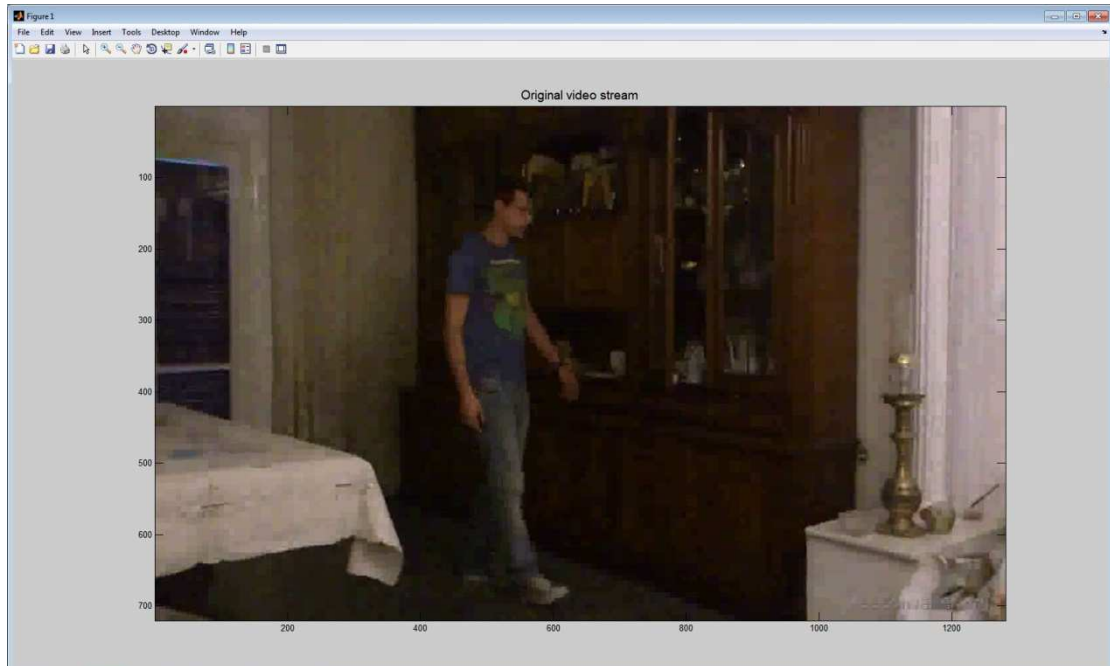
Εικόνα 59: Στιγμιότυπο 2 από την αναπαραγωγή βίντεο 5



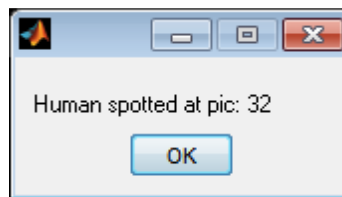
Εικόνα 60: Στιγμιότυπο 3 από την αναπαραγωγή βίντεο 5



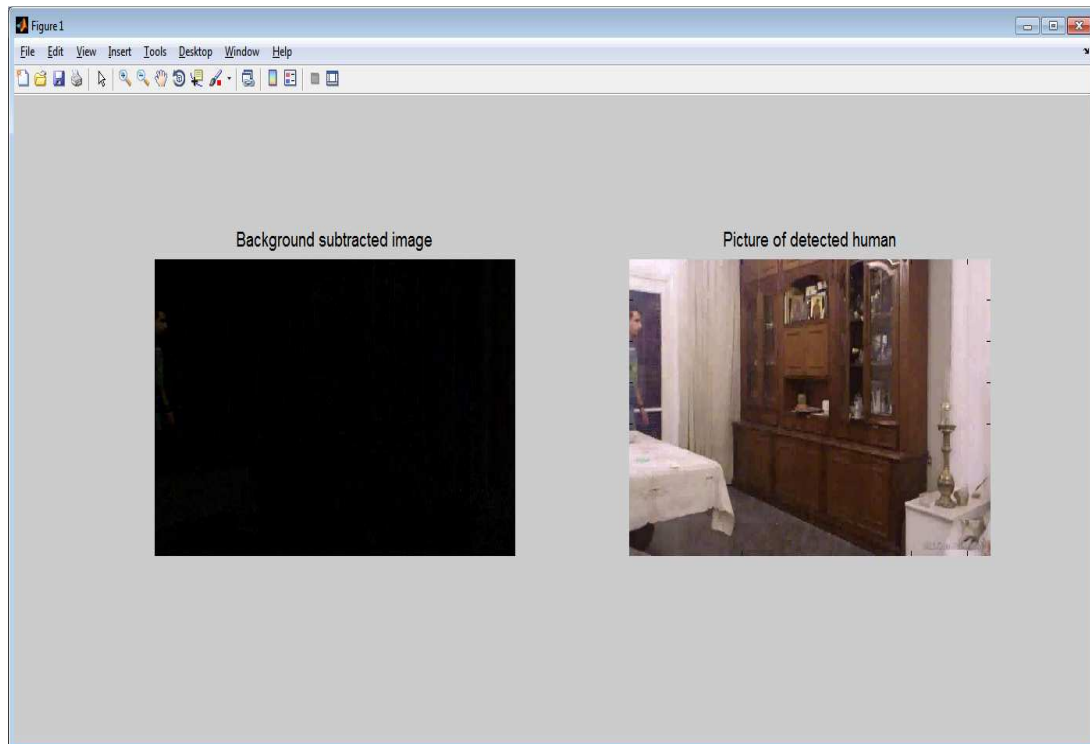
Εικόνα 61: Στιγμιότυπο 4 από την αναπαραγωγή βίντεο 5



Εικόνα 62: Στιγμιότυπο 4 από την αναπαραγωγή βίντεο 5

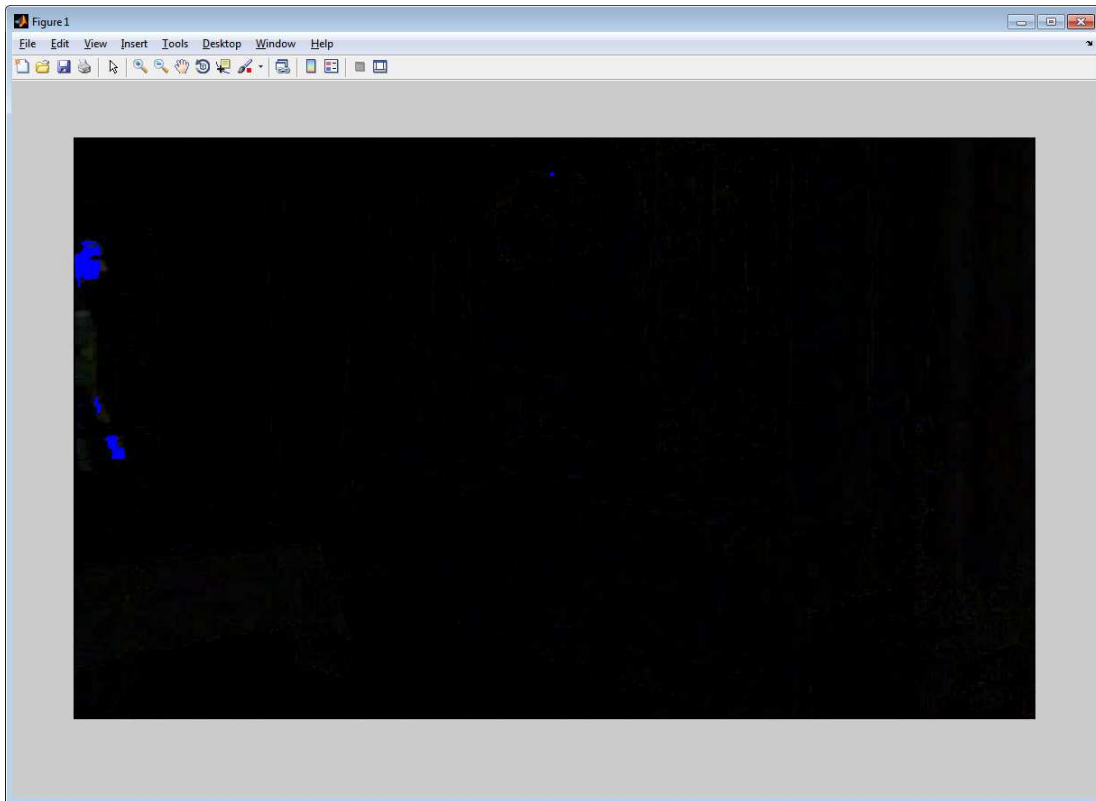


Εικόνα 63: Μήνυμα ανίχνευσης ανθρώπου στο βίντεο 5



Εικόνα 64: Η εικόνα που εντοπίστηκε ανθρώπινη παρουσία στο βίντεο 5 με και χωρίς αφαίρεση background

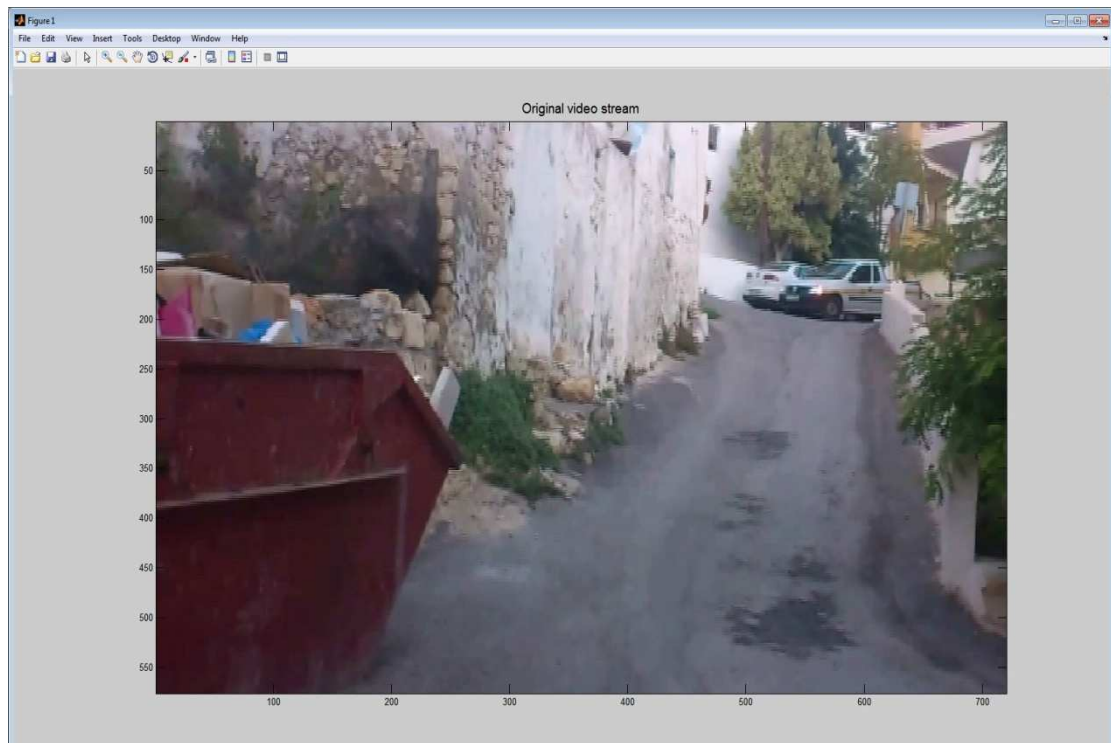
Βλέπουμε ότι η αφαίρεση του background είναι αρκετά αποτελεσματική, παρόλες τις εναλλαγές του φωτισμού και ο αλγόριθμος εντοπισμού pixel στο χρώμα του δέρματος είναι αρκετά ικανοποιητικός βοηθώντας τις να φτάσουμε σε ορθή ανίχνευση ανθρώπου.



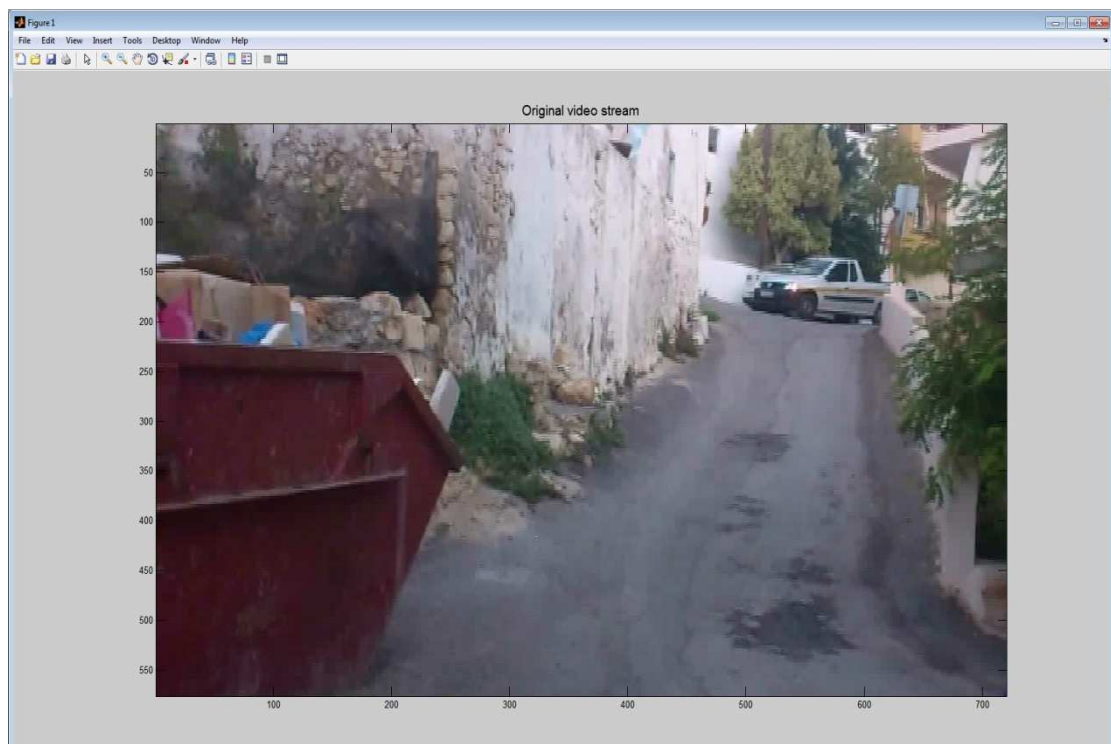
Εικόνα 65: Εντοπισμός pixel στο χρώμα του δέρματος

Στα παραπάνω βίντεο είδαμε την αποδοτικότητα του αλγορίθμου σε εσωτερικούς χώρους. Παρατηρήσαμε ότι καταφέραμε να εντοπίσουμε ανθρώπινη παρουσία, πολύ αποτελεσματικά, σε τις τις περιπτώσεις. Θα συνεχίσουμε δοκιμάζοντας την εφαρμογή σε εξωτερικούς χώρους, σε ένα λιγότερο ελεγχόμενο περιβάλλον με αρκετούς αντιπερισπασμούς.

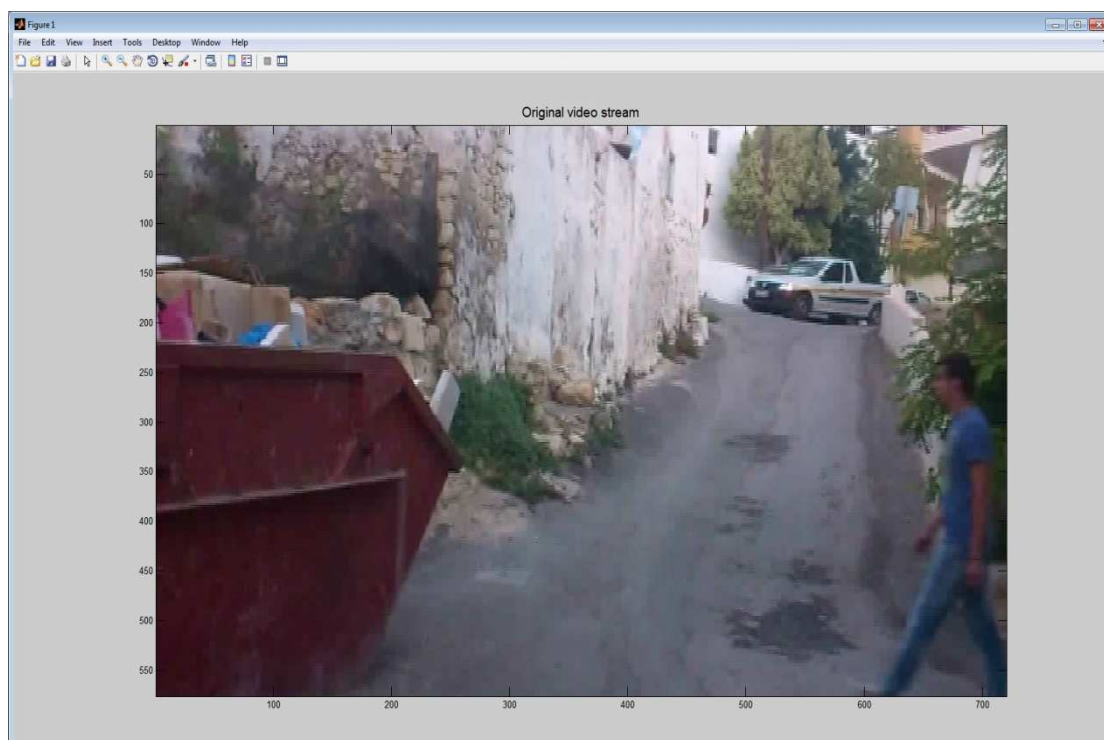
Στη συγκεκριμένη σκηνή βλέπουμε ότι, εκτός από την κίνηση του ανθρώπου έχουμε και την κίνηση διερχόμενου οχήματος. Ωστόσο, η εφαρμογή καταφέρνει να προσπεράσει τη μετακίνηση του οχήματος χωρίς να φτάσει σε εσφαλμένη αναγνώριση.



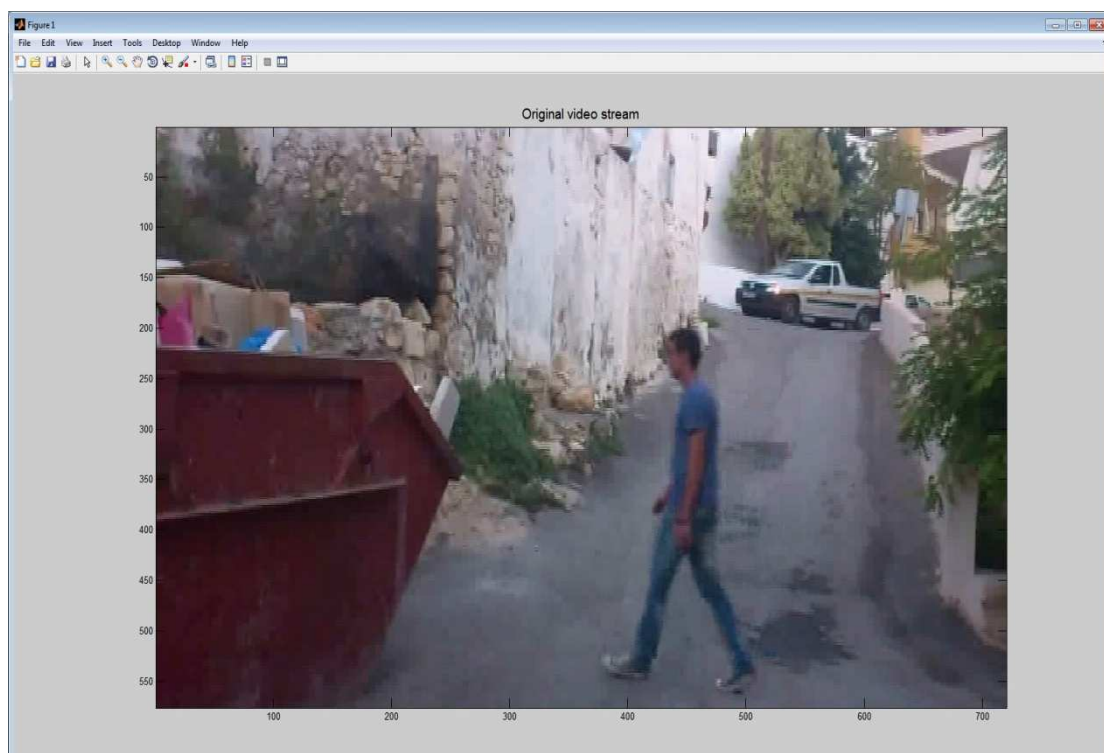
Εικόνα 66: Στιγμιότυπο 1 από την αναπαραγωγή βίντεο 6



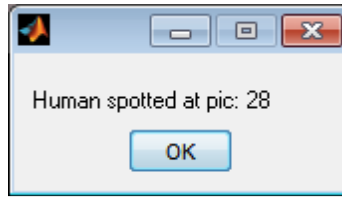
Εικόνα 67: Στιγμιότυπο 2 από την αναπαραγωγή βίντεο 6



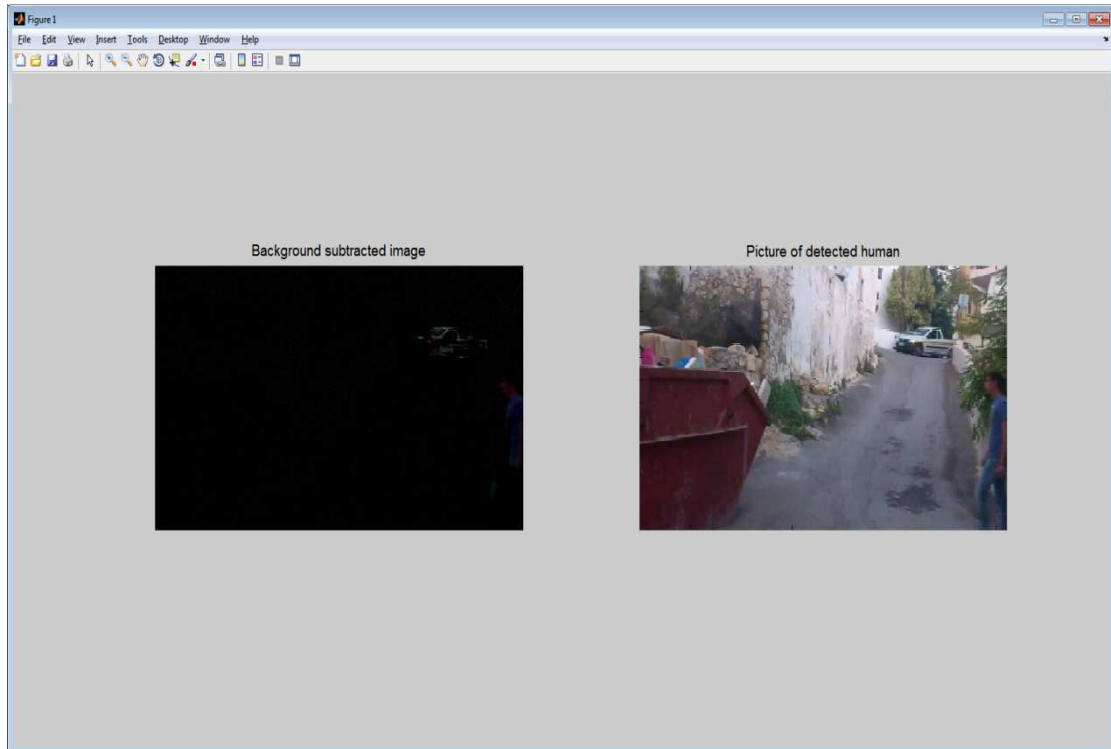
Εικόνα 68: Στιγμιότυπο 3 από την αναπαραγωγή βίντεο 6



Εικόνα 69: Στιγμιότυπο 4 από την αναπαραγωγή βίντεο 6

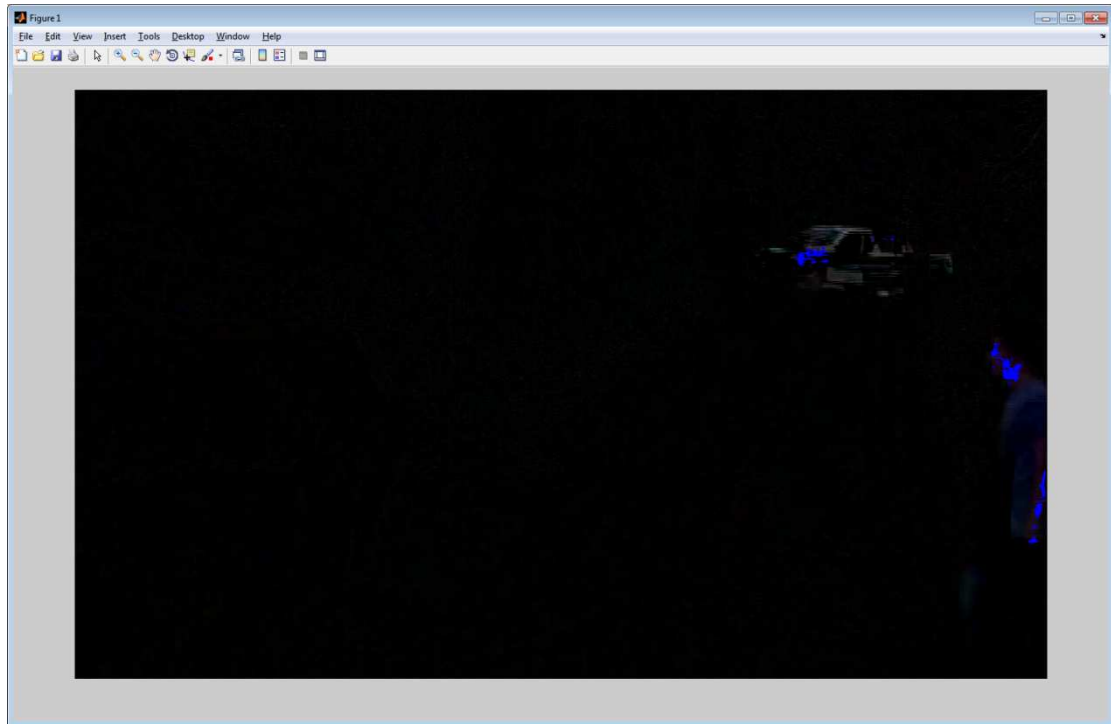


Εικόνα 70: Μήνυμα ανίχνευσης ανθρώπου στο βίντεο 6



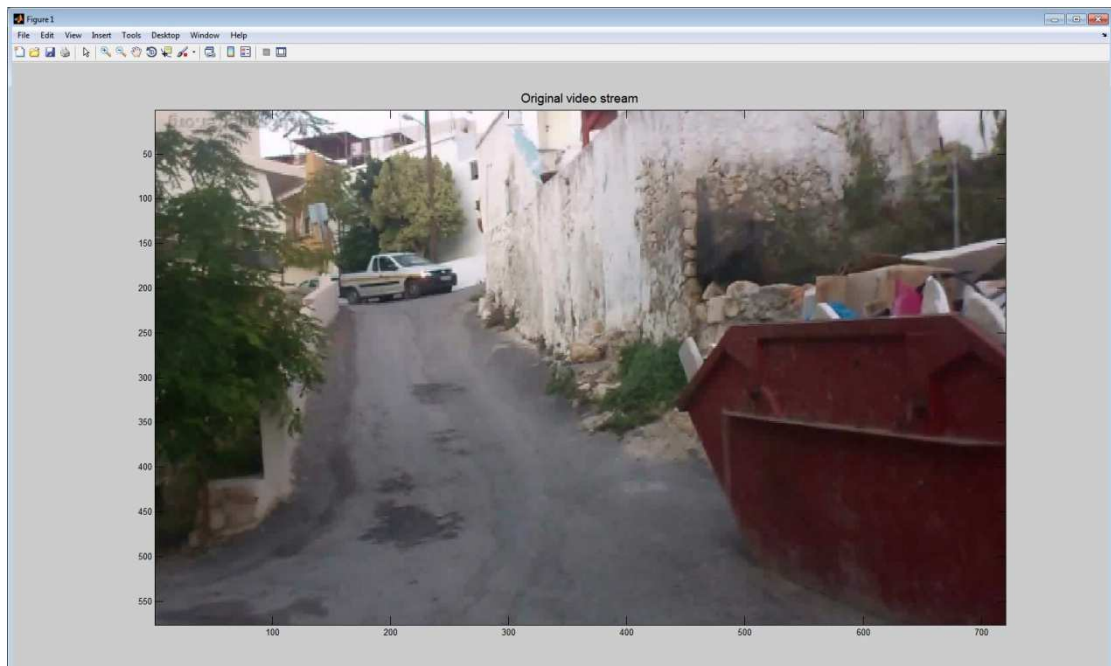
Εικόνα 71: Η εικόνα που εντοπίστηκε ανθρώπινη παρουσία στο βίντεο 6 με και χωρίς αφαίρεση background

Βλέπουμε στο συγκεκριμένο βίντεο ότι η εφαρμογή είναι αρκετά ελαστική με την μετακίνηση αντικειμένων που δεν παρουσιάζουν ενδιαφέρον, όπως αυτή των φύλλων των δέντρων και η μετακίνηση του οχήματος αρκεί να μην έχουν χρώμα παρόμοιο με αυτό του ανθρώπινου δέρματος. Στην εικόνα που ακολουθεί βλέπουμε ότι εντοπίζει pixel που είναι στο χρώμα του δέρματος αλλά αυτά δεν είναι αρκετά σε αριθμό για να προκαλέσουν εσφαλμένη ανίχνευση.

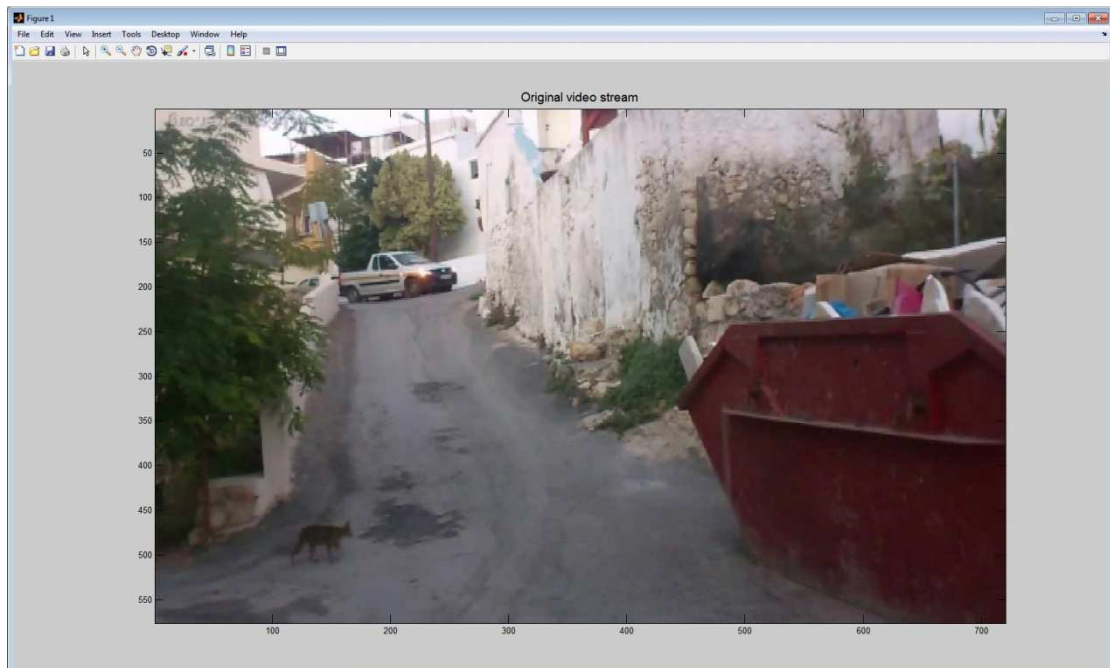


Εικόνα 72: Εντοπισμός pixel στο χρώμα του δέρματος για το βίντεο 6

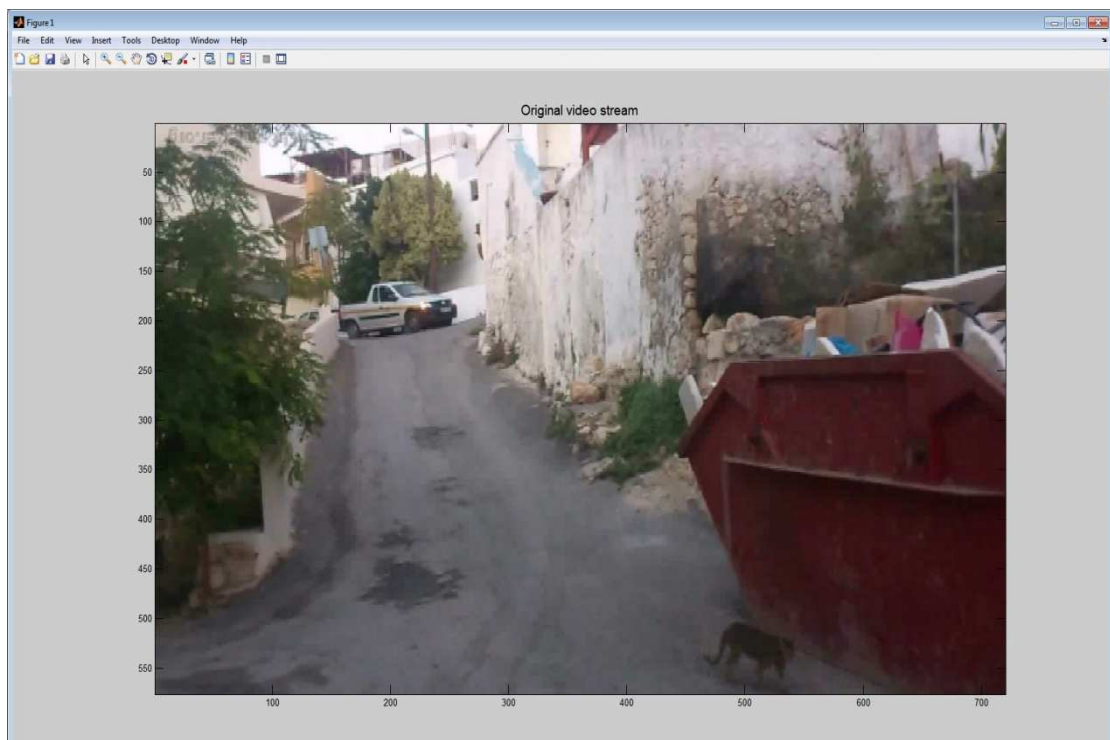
Τελειώνοντας να επισημάνουμε ότι η εφαρμογή δεν επηρεάζεται από μικρές αλλαγές στο foreground(προσκήνιο), όπως είναι κινήσεις που προκαλούνται από μικρά ζώα. Στο παράδειγμα που ακολουθεί έχουμε δυο αλλαγές που προκύπτουν στην εικόνα. Η πρώτη είναι από τα φώτα ενός αυτοκινήτου και η δεύτερη από ένα ζώο που περνάει από το πλάνο.



Εικόνα 73: Στιγμιότυπο 1 από την αναπαραγωγή βίντεο 7

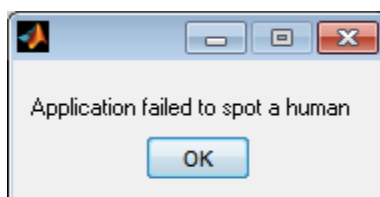


Εικόνα 74: Στιγμιότυπο 2 από την αναπαραγωγή βίντεο 7

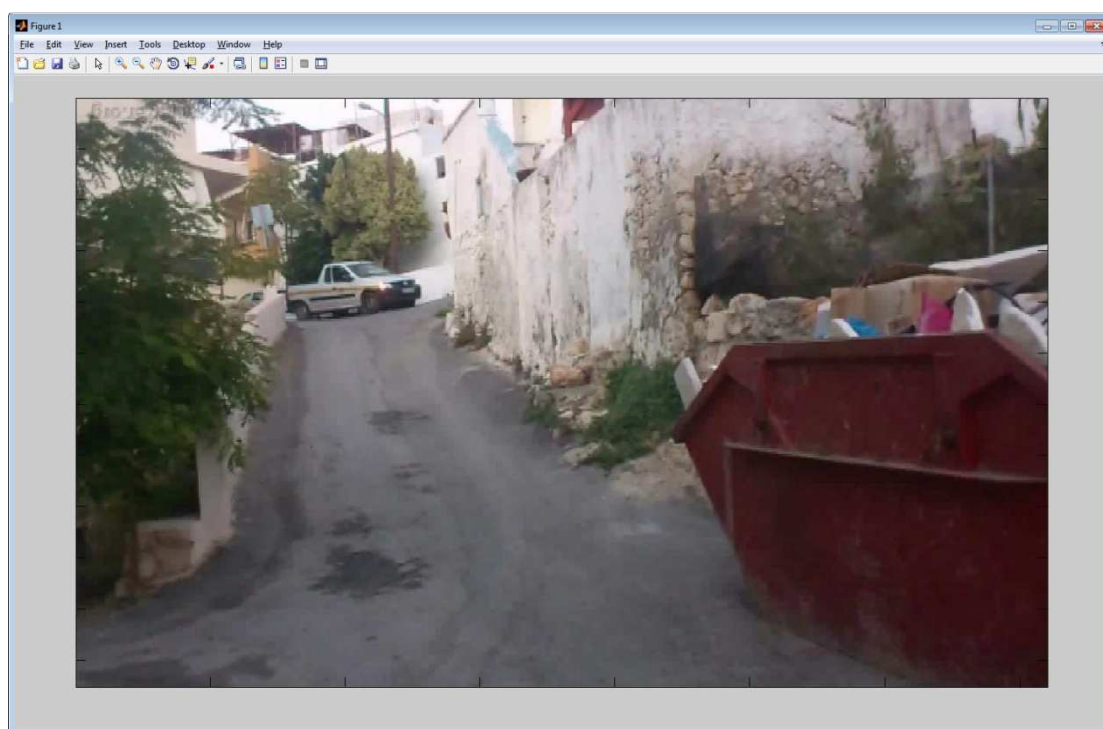


Εικόνα 75: Στιγμιότυπο 3 από την αναπαραγωγή βίντεο 7

Η εφαρμογή δεν εντοπίζει ανθρώπινη παρουσία και εμφανίζει το παρακάτω μήνυμα καθώς και την αρχική εικόνα.



Εικόνα 76: Μήνυμα αντίκρουσης ανθρώπου στο βίντεο 7



Εικόνα 77: Εικόνα επιστροφής εφαρμογής σε περίπτωση που δεν εντοπιστεί άνθρωπος

6 Αποτελέσματα

6.1 Ανασκόπηση

Στην παρούσα πτυχιακή εργασία πραγματοποιήθηκε έρευνα πάνω στις τεχνολογίες και τους τρόπους που χρησιμοποιούνται προκειμένου να επιτευχθεί ανίχνευση ανθρώπινης παρουσίας.

Οι τεχνολογίες που αναλύσαμε εκτενέστερα ήταν η ανίχνευση προσώπου, η ανίχνευση δέρματος, η ανίχνευση κίνησης και τέλος ο εντοπισμός χαρακτηριστικών σημείων.

Έγινε μια περιγραφή του Matlab το οποίο επιλέχθηκε σαν εργαλείο για τη δημιουργία της εφαρμογής μας και τέλος υλοποιήσαμε έναν αλγόριθμο ο οποίος βασίζεται στην αφαίρεση background και στην ανίχνευση δέρματος προκειμένου να εντοπίσει ανθρώπινη παρουσία σε ένα βίντεο. Αναφέραμε τις δυσκολίες που αντιμετωπίσαμε και προτείναμε κάποιες λύσεις για τα προβλήματα αυτά.

6.2 Συμπεράσματα

Το αντικείμενο της πτυχιακής, ήταν μια πρώτη επαφή με τον τομέα της τεχνητής όρασης. Κατανοήσαμε τις δυσκολίες που αντιμετωπίζουν συστήματα τεχνητής νοημοσύνης και συνειδητοποιήσαμε πόσο σημαντικός είναι ο κλάδος αυτός στις μέρες μας.

Σαν άνθρωποι είμαστε ικανοί πολύ εύκολα να αναγνωρίζουμε αντικείμενα που έχουμε ξαναδεί αλλά και να μάθουμε καινούργια αντικείμενα. Μπορούμε να τα αναγνωρίσουμε ανεξάρτητα από το φωτισμό ή την οπτική γωνία που τα κοιτάμε, το ίδιο εύκολα μπορούμε να αναγνωρίσουμε και ανθρώπους σε ένα βίντεο. Αποδείχθηκε ότι δεν είναι τόσο απλό για ένα πρόγραμμα να κάνει το ίδιο πράγμα, καθώς υπάρχουν πολλές μεταβλητές που επηρεάζουν την απόδοσή του.

Καταφέραμε να επιτύχουμε το στόχο μας δημιουργώντας μια απλή εφαρμογή η οποία είναι ικανή να ανιχνεύσει με υψηλό ποσοστό ακρίβειας ανθρώπινη παρουσία σε μερικώς ελεγχόμενα περιβάλλοντα και να διαχειριστεί με επιτυχία κάποιες από τις πιο συνήθεις μεταβολές που μπορούν να εμφανιστούν σε ένα βίντεο.

Βιβλιογραφία

1. A real-world system for human motion detection and tracking. David Moore
California Institute of Technology, [Ηλεκτρονικό]
<http://vision.caltech.edu/~dmoore/dmoore-final-thesis.pdf>
2. HUMAN DETECTION IN VIDEOS 1Muhammad Usman Ghani Khan,
[Ηλεκτρονικό] <http://www.jatit.org>
3. Robust Salient Motion Detection with Complex Background for Real-time Video
Surveillance, Ying-Li Tian and Arun Hampapur, [Ηλεκτρονικό]
[http://www.computer.org/csdl/proceedings/wacv-motion/2005/2271/02/227120030-
abs.html](http://www.computer.org/csdl/proceedings/wacv-motion/2005/2271/02/227120030-abs.html)
4. R. Cutler and L. S. Davis. Robust real-time periodic motion detection, analysis, and
applications, IEEE Transactions on Pattern Analysis and Machine Intelligence
5. A Fusion Approach for Efficient Human Skin Detection, Wei Ren Tan, Chee Seng
Chan, *Member, IEEE*, Pratheepan Yogarajah, and Joan Condell
6. Skin Detection - a Short Tutorial, Ahmed Elgammal, Crystal Muang and Dunxu Hu
Department of Computer Science, Rutgers University, Piscataway, NJ, 08902, USA
7. Night vision, [Ηλεκτρονικό] http://en.wikipedia.org/wiki/Night_vision
8. Night vision, [Ηλεκτρονικό] <http://www.hownightvisionworks.com/#UIC>
9. Face detection, [Ηλεκτρονικό] http://en.wikipedia.org/wiki/Face_detection
10. Human Facial Expression Recognition using Eigen Face and Neural Network,
Pushpaja V. Saudagare, D. S. Chaudhari, International Journal of Engineering and
Advanced Technology (IJEAT), Volume-1, Issue-5, June 2012
11. REAL TIME MOTION DETECTION FOR FAST HUMAN IDENTIFICATION
BASED ON FACE RECOGNITION, Pablo Rivas, Mario I. Chacón, Digital Signal
Processing & Vision Laboratory, Chihuahua Institute of Technology, [Ηλεκτρονικό]
[http://www.academia.edu/966953/Real_time_motion_detection_for_fast_human_ide
ntification_based_on_face_recognition](http://www.academia.edu/966953/Real_time_motion_detection_for_fast_human_identification_based_on_face_recognition)
12. MATLAB, [Ηλεκτρονικό] <http://en.wikipedia.org/wiki/MATLAB>,
http://www.math.uoc.gr/~chatzipa/index_files/matlab/introduction.pdf
13. Y. Song, A perceptual approach to human motion detection and labeling. PhD thesis,
California Institute of Technology, 2003.
14. C. Tomasi and T. Kanade, "Detection and tracking of point features," Tech. Rep.
CMUCS-91-132, Carnegie Mellon University, 1991.